1999–2000
TECHNICAL MANUAL

TABLE OF CONTENTS

CHAPTER 1—BACKGROUND AND OVERVIEW

PURPOSE OF THIS MANUAL

The purpose of this technical manual is to document the technical aspects of the 1999–2000 Maine Educational Assessment (MEA). In the fall of 1999, students in grades 4, 8, and 11 participated in the administration of the revised MEA; during this administration writing, reading, and health education were assessed. In the spring of 2000, students in grades 4, 8, and 11 were administered tests in mathematics, science and technology, social studies, and visual and performing arts. This report provides information about the technical quality of those assessments, including a description of the processes used to develop, administer, and score the tests and to analyze the test results. This report is intended to serve as a guide for replicating and/or improving the procedures in subsequent years.

While some parts of this technical report may be used by educated laypersons, the intended audience is experts in psychometrics and educational research. The report assumes a working knowledge of measurement concepts such as reliability and validity, and statistical concepts such as correlation and central tendency. In some chapters, the reader is presumed to also have basic familiarity with advanced topics in measurement and statistics.

LEARNING RESULTS

Following enactment of the Education Reform Act of 1984, Maine schools undertook a wide variety of initiatives designed to improve the quality of teaching and learning. Many of the lessons learned from those initiatives informed *Maine's Common Core of Learning*, a document published in 1990 that articulates a common vision for education in Maine by defining the knowledge, the skills, and the attitudes that all students should possess upon graduation from high school. In 1993, the legislature directed the state board of education to undertake the next step in education reform by establishing a task force on *Learning Results* that was directed to "develop long-range education goals and standards for school performance and student performance to improve learning results and recommend to the commissioner and to the Legislature a plan for achieving those goals and standards."

After substantial work, the task force presented to the legislature, in January of 1996, a report that contained a series of recommendations together with a set of standards, a plan for implementation, and proposed legislation. After a series of intense hearings during the 1996 legislative session, the legislature adopted much of the work of the task force and directed the department of education and the state board of education to continue to develop the *Learning Results.*

Acting on the recommendations of the task force, the legislature adopted six guiding principles that describe the characteristics of a well-educated person. To fulfill these principles, the legislature required that the department of education and the state board of education develop *Learning Results* within the following eight areas:

Career Preparation
English Language Arts
Health and Physical Education
Mathematics
Modern and Classical Languages
Science and Technology
Social Studies
Visual and Performing Arts

These are not "subjects" in the same sense that we use the word when referring to courses in school. They are areas of learning that will in some cases cut across a number of discrete courses or disciplines.

In response to the legislative directive, the commissioner appointed a working group, known as the Critical Review Committee, to prepare a draft of standards for consideration by the state board of education and by the legislature. The committee met on numerous occasions during the summer and fall of 1996 to produce this revised document, which was approved in May of 1997 by the 118th legislature.

PURPOSES OF THE MEA

The *Learning Results* are just one part of an educational system. As goals for what all students should know and be able to do upon finishing school, they are not written to prescribe a minimum or "passing" standard. The setting of minimum requirements is the function of assessments that are separate from the creation of academic goals.

Because some students are ready for assessment at earlier stages than others, no assumption is made about when a standard might be achieved.

The statute passed in April of 1996 includes the following provisions relating to assessment:

> Student achievement of the learning results…must be measured by a combination of state and local assessments to measure progress and ensure accountability. The 4th-grade, 8th-grade, and 11th-grade results of the Maine Educational Assessment, the "MEA," are the state assessments used to measure achievement of the learning results. The 4th-grade and 8th-grade MEA must be used to measure achievement of the learning results beginning in the 1998–99 school year. Local school administrative units may develop additional assessments to measure achievement of the learning results, including student portfolios, performances, demonstrations, and other records of achievements.

An Assessment Design Team composed of Maine educators and assessment specialists has been established to redesign state-level assessments and to assist in the development of high-quality local assessments that will be used to measure student achievement of the *Learning Results*. The statewide assessment system they are developing will

- align with Maine's *Learning Results;*
- utilize multiple measures of learning;
- ensure fair and equitable assessment for all students;
- utilize recognized, relevant technical standards for assessment;
- provide understandable information to educators, parents, students, the public, and the media;
- provide professional development opportunities for teachers, administrators, and future educators; and
- be practical and manageable.

ORGANIZATION OF THIS MANUAL

The organization of this manual is based on the conceptual flow of an assessment's life span; it begins with the initial test specification and addresses all the intermediate steps that lead to final score reporting. Section I covers the development of the MEA tests. It consists of eight chapters, covering general design issues, the test development process, and the specific designs of the English language arts, mathematics, science and technology, social studies, visual and performing arts, and health education assessments. Section II consists of a single chapter describing the administration of the tests. Section III contains six chapters, covering scoring, equating, item analysis, reliability, validity, and score reporting. Section IV contains references and Section V contains the appendices.

SECTION I: ASSESSMENT DEVELOPMENT

CHAPTER 2—OVERVIEW OF TEST DESIGN

LEARNING RESULTS

MEA questions are directly linked to the **content standards** and the **performance indicators** described in Maine's *Learning Results*. The content standards are the basis for the reporting categories developed for each subject area; the performance indicators are used to help guide the development of test questions. No other content or process is subject to statewide assessment. An item may address part, all, or several of the performance indicators.

ITEM TYPES

Maine's educators and students were familiar with most of the question types that were used in the new assessment program, although one new type—the extended-response question—was used as well. The types of questions used and the functions of each are described below.

Multiple-choice questions were used, in part, to provide breadth of coverage of a subject area. Because they require no more than a minute for most students to answer, these questions make efficient use of limited testing time and allow coverage of a wide range of knowledge and skills.

Short-answer questions were used to assess students' skills and their abilities to work with brief, well-structured problems that had one or a very limited number of solutions (e.g., mathematical computations). Short-answer questions require approximately two to five minutes for most students to answer. The advantage of this type of question is that it requires students to demonstrate knowledge and skills by generating, rather than merely selecting, an answer.

Constructed-response questions are the same as the open-response questions that have been used in past years of the MEA. These questions typically require students to use higher-order thinking skills—evaluation, analysis, summarization, and so on—in constructing a satisfactory response. Constructed-response questions should take most students approximately five to ten minutes to complete. It should be noted that the use of previously released MEA questions to prepare students to answer this kind of question was appropriate and encouraged.

Extended-response questions are a type of question that had not been used previously in the MEA until 1998–99. These questions assess students' ability to analyze and solve challenging problems based on real-world, age-appropriate situations that call for multiple approaches and may have more than one solution. An ability to communicate and justify a solution through the use of writing, tables, charts, and/or graphic displays contributes to a student's success in many of the extended-response questions. This type of question requires approximately ten to twenty minutes for most students to complete.

COMMON-MATRIX DESIGN

In 1999–00, the MEA continued to measure what students know and are able to do by using a greater variety of question types. The tests continued to be structured using both **common** and **matrix-sampled** questions. Common questions are those taken by all students at a given grade level. In addition, a larger pool of matrix-sampled questions is divided among the multiple forms of the test at each grade level. (There were twelve forms of the test in 1999–00.) Each student takes only one form of the test and so answers a fraction of the matrix-sampled questions in the entire pool. This design, which has been used throughout the MEA's history, provides reliable and valid results at the student level. It also provides for a greater breadth of coverage of a subject area for school results while minimizing testing time through the use of matrix-sampled questions.

In 1999–00, the reports continued to only report out common scores in the results for ease of understanding them. If student results were based on common and matrix-sampled questions, one student could score higher than another in raw score, but lower in scaled score. By giving common results only, this type of reversal is avoided.

TEST SESSION TIMES

The MEA tests were given at two different times during the school year: **writing, reading,** and **health education** were administered to all grades in late fall; tests in **all other subject areas** were administered to all grades during a two-week period in early March. Schools were able to schedule testing sessions at any time during the first week of this period, provided they followed the sequence in the scheduling guidelines detailed in test administration manuals. The second week was reserved for makeup testing of students who were absent from initial test sessions.

The timing and scheduling guidelines for MEA tests were based on estimates of the time it would take an average student to respond to each type of question that made up the test:

- multiple-choice questions – 1 minute per question
- short-answer questions – 2 minutes per question
- constructed-response questions – 10 minutes per question
- extended-response questions – 20 minutes per question

For the English language arts reading test, the scheduling guidelines included an estimate of ten minutes to read each passage used in the assessment.

While the guidelines for scheduling were based on the assumption that most students would complete the test within the time estimated, each test session was scheduled so that additional time was provided for students who needed it. One-third additional time was allocated for each session (i.e., sixty-minute sessions were scheduled with an additional twenty minutes; forty-five-minute sessions with an additional fifteen minutes; and thirty-minute sessions with an additional ten minutes).

If additional classroom space was not available for students who required additional time to complete the tests, schools were allowed to consider using another space, such as the guidance office, for this purpose. If additional areas were not available, it was recommended that each classroom being used for test administration be scheduled for the maximum amount of time. Detailed instructions on test administration and scheduling were provided in the *Coordinator's* and *Test Administrator's Manuals*.

CHAPTER 3—TEST DEVELOPMENT PROCESS

DEVELOPMENT COMMITTEE ITEM IDEA GENERATION

The development of the MEA tests continued to be a cooperative effort by committees of Maine teachers, curriculum supervisors, higher education faculty, content specialists of the department of education, and curriculum/assessment specialists employed by the program's contractor, Measured Progress. The committees were structured to represent all areas of the state, and committee members all served rotating terms.

The committees' primary roles were to develop test questions for the MEA and to interpret testing data so that questions could be selected for the program. The MEA development committee for each subject area at grade levels 4, 8, and 11 met several times. In the development phase, the committees reviewed the content standards and test specifications. They also brainstormed or drafted test questions and scoring rubrics to fit those specifications. After the questions were field-tested, the committees reviewed the field-test data and made recommendations about selecting, revising, or eliminating specific questions from the item pool for the operational test. At that time, the committees also confirmed that each question conformed directly to Maine's *Learning Results* and was thus assigned to the appropriate content standard reported in school and district results. Because many MEA questions are released to the public each year, the committees repeat these activities annually as new questions are developed in order to replenish the item pool.

INTERNAL ITEM REVIEW

- The lead or peer test developer within the content specialty reviewed the typed item, the open-response scoring guide, and any reading selections and graphics.
- The content reviewer considered item "integrity"; item content and structure; appropriateness to designated content area; item format; clarity; possible ambiguity; keyability; single "keyness"; appropriateness and quality of reading selections and graphics; and appropriateness of scoring guide descriptions and distinctions (as correlated to the item and within the guide itself).

- The content reviewer also considered scorability and evaluated whether the scoring guide adequately addressed performance on the item.
- Fundamental questions the content reviewer considered, but was not limited to, included the following:
  - What is the item asking?
  - Is the key the only possible key?
  - Is the open-response item scorable as written (are the correct words used to elicit the response defined by the guide)?
  - Is the wording of the scoring guide appropriate and parallel to the item wording?
  - Is the item complete (e.g., with scoring guide, content codes, key, grade level, and contract identified)?
  - Is the item appropriate for the designated grade level?

EXTERNAL ITEM REVIEW

- Item sets were brought to Development Advisory Committee meetings for review and revision.

ITEM EDITING

Editors reviewed and edited the items from the Development Advisory Committee item review to ensure uniform style (based on *The Chicago Manual of Style, 14th Edition*) and adherence to sound testing principles. These principles included the stipulation that items

- were correct with regard to grammar, punctuation, usage, and spelling;
- were written in a clear, concise style;
- contained unambiguous explanations for students as to what was required to attain a maximum score;
- were written at a reading level that would allow the student to demonstrate his or her knowledge of the tested subject matter regardless of reading ability;
- exhibited high technical quality regarding psychometric characteristics;
- had appropriate answer options or score-point descriptors; and
- were free of potentially insensitive content.

REVIEWING AND REFINING

Test developers presented item statistics to the development committees to assist in the committees' recommendation for placement of items into the common and matrix portions of the test. The department of education made the final selections with the assistance of Advanced Systems at a meeting.

OPERATIONAL TEST ASSEMBLY

Test assembly is the sorting and laying out of item sets into test forms. Criteria considered during this process included the following:

- Content coverage/match to test design. The curriculum specialist completed an initial sorting of items into sets based on a balance of content categories across sessions and forms, as well as a match to the test design (e.g., number of multiple-choice, short-answer, and open-response items).
- Item difficulty and complexity. Item statistics drawn from the data analysis of previously tested items were used to ensure that there were similar levels of difficulty and complexity across forms.
- Visual balance. Item sets were reviewed to ensure that each reflected a similar length and "density" of selected items (e.g., length/complexity of reading selections or number of graphics).
- Option balance. Each item set was checked to verify that it contained a roughly equivalent number of key options (As, Bs, Cs, and Ds).
- Name balance. Item sets were reviewed to ensure that a diversity of names was used.
- Bias. Each item set was reviewed to ensure fairness and balance based on gender, ethnicity, religion, socioeconomic status, and other factors.
- Page fit. Item placement was modified to ensure the best fit and arrangement of items on any given page.
- Facing-page issues. For multiple items associated with a single stimulus (a graphic or a reading selection), consideration was given to whether those items needed to begin on a left- or right-hand page, as well as to the nature and the amount of material that needed to be placed on facing pages. These considerations served to minimize the amount of page flipping required of the students.

- Relationships between forms. Sets of common items were placed identically in each version of the forms. Although matrix-sampled item sets differed from form to form, they took up the same number of pages in each form so that sessions and content areas began on the same page in every form. Therefore, the number of pages needed for the longest form often determined the layout of each form.
- Visual appeal. The visual accessibility of each page of the form was always taken into consideration, including such aspects as the amount of white space, the density of the text, and the number of graphics.

EDITING DRAFTS OF OPERATIONAL TESTS

Any changes made by the test construction specialist had to be reviewed and approved by the test developer. Once a form had been laid out in what was considered its final form, it was reread to identify any final considerations, including the following:

- Editorial changes. All text was scrutinized for editorial accuracy, including consistency of instructional language, grammar, spelling, punctuation, and layout. Advanced Systems' publishing standards are based on *The Chicago Manual of Style, 14th Edition*.
- Keying items. Items were reviewed for any information that might "key" or provide information that would help students answer another item. Decisions about moving keying items were based on the severity of the key-in and the placement of the items in relation to each other within the form.
- Key patterns. The final sequence of keys was reviewed to ensure that the order appeared random (i.e., no recognizable pattern and no more than three of the same key in a row).

BRAILLE AND LARGE-PRINT TRANSLATION

Form one for grades 4, 8, and 11 tests was translated into Braille by a subcontractor who specializes in test materials for blind and visually handicapped students. In addition, form one for each grade was adapted into a large-print version.

CHAPTER 4—DESIGN OF THE ENGLISH LANGUAGE ARTS ASSESSMENT

READING

BLUEPRINT

As indicated earlier, the English language arts framework for reading was based on Maine's *Learning Results*, which identifies five **content standards** that apply specifically to reading and reading comprehension. Those content standards are

- **Process of reading (A):** Students use the skills and the strategies of the reading process to comprehend, interpret, evaluate, and appreciate what they have read.
- **Literature and culture (B):** Students use reading, listening, and viewing strategies to experience, understand, and appreciate literature and culture.
- **Language and images (C):** Students demonstrate an understanding of how words and images communicate.
- **Informational texts (D):** Students apply reading, listening, and viewing strategies to informational texts across all areas of curriculum.
- **Research-related writing and speaking (H):** Students work, write, and speak effectively in connection with research in all content areas.

The content standards have been adapted to create a reporting category framework for reading, as shown below.

Passage Type	Comprehension of Literary and Informational Texts			Total
	Reading Comprehension and Literary Analysis	A. Process of Reading	C. Language and Images	
B. Literature and Culture: Literary Passages				50%
D. Informational Texts: Content Passages Practical Passages				50% (30%) (20%)
Total	80%	20%		100%

CONTENT SPECS

The first major reporting category at the student, school, and district levels is "comprehension of literary and informational texts." The data generated for this reporting category was based on questions related to three types of reading passages that reflect standards B and D of the English language arts (ELA) *Learning Results*. The passage types were identical to those that have been used in the MEA in past years. Fifty percent of the passages were literary works; 30 percent were selected from content pieces (see explanation below); and 20 percent were drawn from practical sources (see explanation below).

Passages included both long and short "authentic" texts selected from reading sources that students at each grade level would be likely to encounter in their classroom and in their independent reading. None of the passages were written specifically for the assessment, but instead they were collected from published works.

- **Literary passages** are represented by a variety of genres—modern narratives; diary entries; drama; poetry; biographies; essays; excerpts from novels; short stories; and traditional narratives, such as fables, myths, and folktales.
- **Content passages** are primarily informational and often deal with the areas of science and social studies. They are drawn from such sources as newspapers, magazines, and books.
- **Practical passages** are functional materials that instruct or advise the reader—for example, directions, reference tools, or manuals.

The main difference in the passages used for grades 4, 8, and 11 was their degree of difficulty. All passages were selected to be appropriate for the intended audience; however, the ideas expressed became increasingly more complex at grade levels 8 and 11.

The questions related to these passages required students to demonstrate their skills in both literal comprehension (where the answer is stated explicitly in the text) and inferential comprehension (where the answer is implied by the text and/or the text must be connected to relevant prior knowledge to determine an answer). In addition, some questions focused on the reading skills reflected in content standards A and C of the *Learning Results*. Questions of this type require students to use the skills and strategies of reading to answer questions—for example, how to identify the author's principal purpose, such as to persuade, entertain, or inform—and to demonstrate their understanding of how words and images communicate to readers.

ITEM TYPES

The MEA English language arts assessment in reading included multiple-choice, short-answer, constructed-response, and extended-response questions. Short-answer questions, which were new in the revised MEA, required students to write an answer consisting of several phrases or short sentences. Each type of question was worth a specific number of points in the student's total language arts score, as shown below.

Type of Question	Possible Score Points
Multiple-Choice	0–1
Short-Answer	0–2
Constructed-Response	0–4
Extended-Response	0–8

TEST DESIGN

The table below summarizes the numbers and types of questions that were used in the MEA reading assessment for 1999–00.

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
2A	6	2	1						25
2B	6	2	2						25
3A	6	1		1					45
3B					6	2	1		25

Key

- MC = multiple-choice questions
- SA = short-answer questions
- CR = constructed-response questions
- ER = extended-response questions

The charts on the following pages outline the total number of possible points—as reported—by learning results and item type.

ENGLISH LANGUAGE ARTS—READING
Number of Points Possible
Grade 4

Standard and Passage	COMMON						MATRIX					TOTAL POSSIBLE POINTS
Reading Process and Language (Standards A and C)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**Points**	**Percent**	**54**
Choosing a Companion	1	0	0	0	1	100	0	0	0	0	0	1
Sarah, Plain and Tall	1	0	0	0	1	100	0	0	0	0	0	1
The Fastest Kid in Fifth Grade	1	2	0	0	3	100	0	0	0	0	0	3
Alcove Spring	0	0	0	0	0	0	2	0	0	0	0	2
Amazing Spiders	0	0	0	0	0	0	1	0	0	1	100	1
Bacon-Tomato Sandwiches	0	0	0	0	0	0	1	0	4	5	100	5
Be a Junk Food Detective	0	0	0	0	0	0	1	0	0	1	100	1
Brian's Winter	0	0	0	0	0	0	1	0	0	1	100	1
Cleaning Up the Ocean	0	0	0	0	0	0	1	0	0	1	100	1
Drinking Milk is Good for Birds	0	0	0	0	0	0	2	2	4	8	100	8
Home Grown Hydras	0	0	0	0	0	0	1	0	0	1	100	1
I'm Going To Be Famous	0	0	0	0	0	0	1	0	0	1	100	1
Let's Write a True Life Story	0	0	0	0	0	0	1	0	4	5	100	5
Marsha	0	0	0	0	0	0	2	0	0	2	100	2
On My Own With Alex	0	0	0	0	0	0	3	2	0	5	100	5
One Brave Summer	0	0	0	0	0	0	0	0	0	0	0	0
Rats Don't	0	0	0	0	0	0	3	0	0	3	100	3
Ruby	0	0	0	0	0	0	2	0	0	2	100	2
Secret Place	0	0	0	0	0	0	2	0	0	2	100	2
Spring Poems	0	0	0	0	0	0	2	2	0	4	100	4
Welcome to the Inventors Club	0	0	0	0	0	0	3	2	0	5	100	5
Reading Comprehension (Standards B and D)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**Points**	**Percent**	**162**
Choosing a Companion	5	4	4	0	13	100	0	0	0	0	0	13
Sarah, Plain and Tall	5	2	0	8	15	100	0	0	0	0	0	15
Shivers and Goose Bumps	3	2	4	0	9	100	0	0	0	0	0	9
The Fastest Kid in the Fifth Grade	2	0	4	0	6	100	0	0	0	0	0	6
Alcove Spring	0	0	0	0	0	0	4	4	4	12	100	12
Amazing Spiders	0	0	0	0	0	0	2	2	0	4	100	4
Bacon-Tomato Sandwiches	0	0	0	0	0	0	2	2	0	4	100	4
Be a Junk Food Detective	0	0	0	0	0	0	2	2	0	4	100	4

Brian's Winter	0	0	0	0	0	0	2	2	0	4	100	4
Cleaning Up the Ocean	0	0	0	0	0	0	2	2	0	4	100	4
Drinking Milk is Good for Birds	0	0	0	0	0	0	4	2	0	6	100	6
Home Grown Hydras	0	0	0	0	0	0	2	2	4	8	100	8
I'm Going to be Famous	0	0	0	0	0	0	5	4	4	13	100	13
Let's Write a True Life Story	0	0	0	0	0	0	2	2	0	4	100	4
Marsha	0	0	0	0	0	0	1	2	4	7	100	7
On My Own With Alex	0	0	0	0	0	0	3	2	4	9	100	9
One Brave Summer	0	0	0	0	0	0	3	2	4	9	100	9
Rats Don't	0	0	0	0	0	0	3	4	4	11	100	11
Ruby	0	0	0	0	0	0	1	2	0	3	100	3
Secret Place	0	0	0	0	0	0	1	2	0	3	100	3
Spring Poems	0	0	0	0	0	0	1	0	4	5	100	5
Welcome to the Inventors Club	0	0	0	0	0	0	3	2	4	9	100	9
Literature & Culture (Standard B)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**Points**	**Percent**	**88**
Sarah, Plain and Tall	5	2	0	8	15	100	0	0	0	0	0	15
Shivers and Goose Bumps	2	0	4	0	6	100	0	0	0	0	0	6
Alcove Spring	0	0	0	0	0	0	4	4	4	12	100	12
Brian's Winter	0	0	0	0	0	0	2	2	0	4	100	4
I'm Going to be Famous	0	0	0	0	0	0	5	4	4	13	100	13
Marsha	0	0	0	0	0	0	1	2	4	7	100	7
One Brave Summer	0	0	0	0	0	0	3	2	4	9	100	9
Rats Don't	0	0	0	0	0	0	3	4	4	11	100	11
Ruby	0	0	0	0	0	0	1	2	0	3	100	3
Secret Place	0	0	0	0	0	0	1	2	0	3	100	3
Spring Poems	0	0	0	0	0	0	1	0	4	5	100	5
Informational Texts (Standard D)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**Points**	**Percent**	**74**
Choosing a Companion	5	4	4	0	13	100	0	0	0	0	0	13
Shivers and Goose Bumps	3	2	4	0	9	100	0	0	0	0	0	9
Amazing Spiders	0	0	0	0	0	0	2	2	0	4	100	4
Bacon-Tomato Sandwiches	0	0	0	0	0	0	2	2	0	4	100	4
Be a Junk Food Detective	0	0	0	0	0	0	2	2	0	4	100	4
Cleaning Up the Ocean	0	0	0	0	0	0	2	2	0	4	100	4
Drinking Milk is Good for Birds	0	0	0	0	0	0	4	2	0	6	100	6
Home Grown Hydras	0	0	0	0	0	0	2	2	4	8	100	8
Let's Write a True Life Story	0	0	0	0	0	0	2	2	0	4	100	4
On My Own With Alex	0	0	0	0	0	0	3	2	4	9	100	9
Welcome to the Inventors Club	0	0	0	0	0	0	3	2	4	9	100	9

ENGLISH LANGUAGE ARTS—READING
Number of Points Possible
Grade 8

Standard and Passage	COMMON						MATRIX						TOTAL POSSIBLE POINTS
Reading Process and Language (Standards A and C)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**50**
Leave Well Enough Alone	1	0	4	0	5	100	0	0	0	0	0	0	5
On Being a Granddaughter	2	0	0	0	2	100	0	0	0	0	0	0	2
The Ant Parade	2	0	4	0	6	100	0	0	0	0	0	0	6
An American Childhood	0	0	0	0	0	0	2	0	0	0	2	100	2
Cool Science- A Lesson Runs Through It	0	0	0	0	0	0	1	0	0	0	1	100	1
Diary of Anne Frank/ Zlata's Diary	0	0	0	0	0	0	1	0	0	0	1	100	1
Go Fly a Kite	0	0	0	0	0	0	1	0	0	0	1	100	1
Principles of Art	0	0	0	0	0	0	1	0	0	0	1	100	1
Right Smart O'Wind	0	0	0	0	0	0	6	0	4	0	10	100	10
Road Runner	0	0	0	0	0	0	2	0	0	0	2	100	2
The Base Stealer	0	0	0	0	0	0	2	0	0	0	2	100	2
The Debate Over Closing the Door to America	0	0	0	0	0	0	2	0	0	0	2	100	2
The Life of the Ladybird Beetle	0	0	0	0	0	0	1	0	4	0	5	100	5
Uncle Joe	0	0	0	0	0	0	2	2	0	0	4	100	4
Why I Never Shoot Bears	0	0	0	0	0	0	1	0	0	0	1	100	1
Wreck of the Monkey Cage	0	0	0	0	0	0	2	0	0	0	2	100	2
You Can Be An Inventor	0	0	0	0	0	0	3	0	0	0	3	100	3
Reading Comprehension (Standards B and D)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**162**
Leave Well Enough Alone	2	2	0	0	4	100	0	0	0	0	0	0	4
On Being a Granddaughter	4	2	0	8	14	100	0	0	0	0	0	0	14
The Ant Parade	1	2	0	0	3	0	0	0	0	0	0	0	3
Turf Tickers	6	4	4	0	14	100	0	0	0	0	0	0	14
An American Childhood	0	0	0	0	0	0	4	4	4	0	12	100	12
Cool Science—A Lesson Runs Through It	0	0	0	0	0	0	2	2	0	0	4	100	4
Children of the River	0	0	0	0	0	0	3	2	0	0	5	100	5

Diary of Anne Frank/Zlata's Diary	0	0	0	0	0	0	2	2	0	0	4	100	4
Gentle Friends, Essential Allies…	0	0	0	0	0	0	3	2	0	0	5	100	5
Go Fly a Kite	0	0	0	0	0	0	5	4	4	0	13	100	13
Graduation Morning	0	0	0	0	0	0	3	2	0	0	5	100	5
Niagara Falls	0	0	0	0	0	0	3	2	4	0	9	100	9
Principles of Art	0	0	0	0	0	0	2	2	4	0	4	100	4
Right Smart O'Wind	0	0	0	0	0	0	0	4	0	0	4	100	4
Road Runner	0	0	0	0	0	0	1	2	4	0	7	100	7
The Base Stealer	0	0	0	0	0	0	1	2	4	0	7	100	7
The Debate Over Closing the Door to America	0	0	0	0	0	0	4	4	4	0	12	100	12
The Life of the Ladybird Beetle	0	0	0	0	0	0	5	4	0	0	9	100	9
Uncle Joe	0	0	0	0	0	0	4	2	4	0	10	100	10
Why I Never Shoot Bears	0	0	0	0	0	0	2	2	4	0	8	100	8
Wreck of the Monkey Cage	0	0	0	0	0	0	1	2	4	0	7	100	7
You Can Be An Inventor	0	0	0	0	0	0	0	2	0	0	2	100	2
Literature and Culture (Standard B)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**78**
Leave Well Enough Alone	2	2	0	0	4	100	0	0	0	0	0	0	4
On Being a Granddaughter	4	2	0	8	14	100	0	0	0	0	0	0	14
An American Childhood	0	0	0	0	0	0	4	4	4	0	12	100	12
Children of the River	0	0	0	0	0	0	3	2	0	0	5	100	5
Diary of Anne Frank/Zlata's Diary	0	0	0	0	0	0	2	2	0	0	4	100	4
Graduation Morning	0	0	0	0	0	0	3	2	0	0	5	100	5
Right Smart O'Wind	0	0	0	0	0	0	0	2	0	0	2	100	2
The Base Stealer	0	0	0	0	0	0	1	2	4	0	7	100	7
Uncle Joe	0	0	0	0	0	0	4	2	4	0	10	100	10
Why I Never Shoot Bears	0	0	0	0	0	0	2	2	4	0	8	100	8
Wreck of the Monkey Cage	0	0	0	0	0	0	1	2	4	0	7	100	7
Informational Texts (Standard D)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**		**Points**	**Percent**	**88**
The Ant Parade	1	2	0	0	3	100	0	0	0	0	0	0	3
Turf Tickers	6	4	4	0	14	100	0	0	0	0	0	0	14
Cool Science—A Lesson Runs Through It	0	0	0	0	0	0	2	2	0	0	4	100	4
Gentle Friends, Essential Allies…	0	0	0	0	0	0	3	2	0	0	5	100	5
Go Fly A Kite	0	0	0	0	0	0	5	4	4	0	13	100	13
Niagara Falls	0	0	0	0	0	0	3	2	4	0	9	100	9
Principles of Art	0	0	0	0	0	0	2	2	4	0	8	100	8
Right Smart O'Wind	0	0	0	0	0	0	0	2	0	0	2	100	2

Road Runner	0	0	0	0	0	0	1	2	4	0	7	100	7
The Debate Over Closing the Door to America	0	0	0	0	0	0	4	4	4	0	12	100	12
The Life of the Ladybird Beetle	0	0	0	0	0	0	5	4	0	0	9	100	9
You Can Be An Inventor	0	0	0	0	0	0	0	2	0	0	2	100	2

ENGLISH LANGUAGE ARTS—READING
Number of Points Possible
Grade 11

Standard and Passage	COMMON						MATRIX						TOTAL POSSIBLE POINTS
Reading Process and Language (Standards A and C)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**64**
A Father Sees A Son Nearing Manhood	0	0	0	8	8	100	0	0	0	0	0	0	8
Luncinda Matlock	2	0	0	0	2	100	0	0	0	0	0	0	2
Rosa Parks Describes Her Arrest	1	2	0	0	3	100	0	0	0	0	0	0	3
Winterizing Lawn Mowers	1	0	4	0	5	100	0	0	0	0	0	0	5
A Day At The Theatre	0	0	0	0	0	0	2	2	4	0	8	100	8
A Presidential Candidate	0	0	0	0	0	0	2	0	4	0	6	100	6
At Harvesttime	0	0	0	0	0	0	1	0	0	0	1	100	1
Chief Joseph Of The Nez Perce Speaks	0	0	0	0	0	0	1	0	0	0	1	100	1
Children Of The Sun	0	0	0	0	0	0	1	0	0	0	1	100	1
Dead Snails Leave No Trails	0	0	0	0	0	0	3	0	0	0	3	100	3
Deer Among Cattle	0	0	0	0	0	0	2	0	0	0	2	100	2
Discover White Water Rafting	0	0	0	0	0	0	1	0	0	0	1	100	1
I Wandered Lonely As A Cloud	0	0	0	0	0	0	1	0	0	0	1	100	1
Life In The Thirteen Colonies	0	0	0	0	0	0	1	0	0	0	1	100	1
Mt. Katahdin Via The Knife Edge	0	0	0	0	0	0	2	0	0	0	2	100	2
Nearer	0	0	0	0	0	0	2	0	0	0	2	100	2
New Directions	0	0	0	0	0	0	2	2	0	0	4	100	4
Piltdown Man	0	0	0	0	0	0	1	0	0	0	1	100	1
Prevent Repetitive Strain At The Keyboard	0	0	0	0	0	0	1	2	0	0	3	100	3
Sweet Season	0	0	0	0	0	0	3	2	0	0	5	100	5
The House on Mango Street	0	0	0	0	0	0	2	0	0	0	2	100	2
William	0	0	0	0	0	0	2	0	0	0	2	100	2
Reading Comprehension (Standards B and D)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**152**
A Father Sees a Son Nearing Manhood	6	2	0	0	8	100	0	0	0	0	0	0	8
Lucinda Matlock	1	2	4	0	7	100	0	0	0	0	0	0	7
Rosa Parks Describes Her Arrest	2	0	4	0	6	100	0	0	0	0	0	0	6
Winterizing Lawn Mowers	5	4	0	0	9	100	0	0	0	0	0	0	9
A Day At The Theater	0	0	0	0	0	0	4	4	4	0	12	100	12
A Presidential Candidate	0	0	0	0	0	0	4	2	0	0	6	100	6
At Harvesttime	0	0	0	0	0	0	2	2	0	0	4	100	4

Chief Joseph Of The Nez Perce Speaks	0	0	0	0	0	0	2	2	4	0	8	100	8
Children Of The Sun	0	0	0	0	0	0	2	2	0	0	4	100	4
Dead Snails Leave No Trails	0	0	0	0	0	0	3	4	4	0	11	100	11
Deer Among Cattle	0	0	0	0	0	0	1	2	4	0	7	100	7
Discover White Water Rafting	0	0	0	0	0	0	2	2	0	0	4	100	4
I Wandered Lonely As A Cloud	0	0	0	0	0	0	2	2	0	0	4	100	4
Life In The Thirteen Colonies	0	0	0	0	0	0	2	2	0	0	4	100	4
Mt. Katahdin Via The Knife Edge	0	0	0	0	0	0	1	2	4	0	7	100	7
Nearer	0	0	0	0	0	0	1	2	4	0	7	100	7
New Directions	0	0	0	0	0	0	4	2	4	0	10	100	10
Piltdown Man	0	0	0	0	0	0	2	2	0	0	4	100	4
Prevent Repetitive Strain At the Keyboard	0	0	0	0	0	0	2	0	4	0	6	100	6
Sweet Season	0	0	0	0	0	0	3	2	4	0	9	100	9
The House on Mango Street	0	0	0	0	0	0	1	2	0	0	3	100	3
William	0	0	0	0	0	0	4	4	4	0	12	100	12
Literature & Culture (Standard B)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**72**
A Father Sees a Son Nearing Manhood	6	2	0	0	8	100	0	0	0	0	0	0	8
Lucinda Matlock	1	2	4	0	7	100	0	0	0	0	0	0	7
A Presidential Candidate	0	0	0	0	0	0	4	2	0	0	6	100	6
At Harvesttime	0	0	0	0	0	0	2	2	0	0	4	100	4
Children Of The Sun	0	0	0	0	0	0	2	2	0	0	4	100	4
Deer Among Cattle	0	0	0	0	0	0	1	2	4	0	7	100	7
I Wondered Lonely As A Cloud	0	0	0	0	0	0	2	2	0	0	4	100	4
Nearer	0	0	0	0	0	0	1	2	4	0	7	100	7
New Directions	0	0	0	0	0	0	4	2	4	0	10	100	10
The House on Mango Street	0	0	0	0	0	0	1	2	0	0	3	100	3
William	0	0	0	0	0	0	4	4	4	0	12	100	12
Informational Texts (Standard D)	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**MC**	**SA**	**CR**	**ER**	**Points**	**Percent**	**80**
Rosa Parks Describes Her Arrest	2	0	4	0	6	100	0	0	0	0	0	0	6
Winterizing Lawn Mowers	5	4	0	0	9	100	0	0	0	0	0	0	9
A Day At The Theater	0	0	0	0	0	0	4	4	4	0	12	100	12
Chief Joseph Of The Nez Perce Speaks	0	0	0	0	0	0	2	2	4	0	8	100	8
Dead Snails Leave No Trails	0	0	0	0	0	0	3	4	4	0	11	100	11
Discover White Water Rafting	0	0	0	0	0	0	2	2	0	0	4	100	4
Life In The Thirteen Colonies	0	0	0	0	0	0	2	2	0	0	4	100	4
Mt. Katahdin Via The Knife Edge	0	0	0	0	0	0	1	2	4	0	7	100	7
Piltdown Man	0	0	0	0	0	0	2	2	0	0	4	100	4
Prevent Repetitive Strain At The Keyboard	0	0	0	0	0	0	2	0	4	0	6	100	6
Sweet Season	0	0	0	0	0	0	3	2	4	0	9	100	9

WRITING

BLUEPRINT

The MEA assessed students' writing skills directly through the use of writing prompts, or topics, to which students responded. Maine's *Learning Results* includes two **content standards** that apply specifically to writing. Those content standards are

- **Standard English conventions (F):** Students write and speak correctly, using conventions of standard written and spoken English.
- **Stylistic and rhetorical aspects of writing and speaking (G):** Students use stylistic and rhetorical aspects of writing and speaking to explore ideas, to present lines of thought, to represent and reflect on human experience, and to communicate feelings, knowledge, and opinions.

Note: Standard E, processes of writing and speaking, addresses students' abilities to use the skills and strategies of the writing process. This standard was assessed at the local level only.

The *Learning Results* standards were adapted to create reporting categories for writing, as shown below.

Stylistic and Rhetorical Aspects of Writing (G)	▪ Idea/topic development ▪ Organization ▪ Supporting detail
Standard English Conventions (F)	▪ Grammar ▪ Spelling ▪ Punctuation ▪ Capitalization ▪ Sentence structure

CONTENT SPECIFICATIONS

Four broad types, or modes, of writing were used in the MEA, as listed below[1]:

- **Narration:** Narrative writing answers the question "What happened?" It tells a story through a sequence of events, so that the reader understands the action.
- **Exposition:** Expository writing informs the reader about something. Methods of exposition include comparison and contrast, illustration, classification, definition, and analysis. Methods of exposition are often combined to accomplish a specific purpose for writing.
- **Description:** Descriptive writing presents the qualities of objects, persons, conditions, and actions.
- **Persuasion/argument:** Persuasive writing uses emotional appeals to bring about a change of attitude, point of view, or feeling. Argumentative writing uses logic and reason to bring about a change of attitude, point of view, or feeling; it shows that a conclusion merits belief because of credible data, evidence, and so on.

The student's audience and purpose for writing also influence the development, the style, and the tone of a written composition. These were specified as part of the prompts and varied by grade level.

GRADE 4: Writing prompts required work in the narrative, expository, and descriptive modes. Formats included letters and essays. The audience for writing included one or more of the following: friends, characters from books, or the reader himself or herself. The purposes for writing were mainly to inform; to describe people, places, or things; and to tell personal stories.

[1] Descriptions are adapted from *Modern Rhetoric*, by Cleanth Brooks and Robert Penn Warren.

GRADE 8: Writing prompts required students to write in the narrative, expository, descriptive, and persuasive modes. Formats included letters, speeches, newspaper articles, and essays. The audience for writing included one or more of the following: classmates, friends, new students, and the reader herself or himself. The purposes for writing were mainly to deliver useful information and to relate personal descriptions or experiences.

GRADE 11: Writing prompts were drawn from the narrative, expository, descriptive, and persuasive modes. Among the formats included were speeches, letters, newspaper articles, and essays. The audience for writing included one or more of the following: friends, classes, potential employers, a school board or other official agency, and the reader herself or himself. The purposes for writing included one or more of the following: applying for a job, persuading someone to read a book, responding to a quotation, or defending an opinion.

In addition, the prompts were developed with the following criteria as guidelines:

- The prompts must be interesting to students.
- The prompts must be accessible to all students (i.e., all students would have something to say about the topic).
- The prompts must generate sufficient text to be effectively scored.

TEST DESIGN

Each student responded to one common writing prompt. The common prompt, administered to all students, elicited narrative writing in the 1999–00 MEA administration. Each student also responded to a common extended-response question that was scored for both reading and writing.

The charts on the following page outline the total number of possible points—as reported—by learning results and item type.

ENGLISH LANGUAGE ARTS—WRITING

Number of Points Possible
Grade 4

Standard	Common Prompt	Percent	Matrix Prompt	Percent	Total Possible Points
Standard English Conventions (Standard F)	8	66	4	34	12
Stylistic and Rhetorical Aspects of Writing (Standard G)	12	67	6	33	18

Number of Points Possible
Grade 8

Standard	Common Prompt	Percent	Matrix Prompt	Percent	Total Possible Points
Standard English Conventions (Standard F)	8	66	4	34	12
Stylistic and Rhetorical Aspects of Writing (Standard G)	12	67	6	33	18

Number of Points Possible
Grade 11

Standard	Common Prompt	Percent	Matrix Prompt	Percent	Total Possible Points
Standard English Conventions (Standard F)	8	66	4	34	12
Stylistic and Rhetorical Aspects of Writing (Standard G)	12	67	6	33	18

CHAPTER 5—DESIGN OF THE MATHEMATICS ASSESSMENT

BLUEPRINT

The mathematics framework was based on Maine's *Learning Results*, which identifies eleven **content standards**, as shown below:

- **Numbers and number sense (A):** Students understand and demonstrate a sense of what numbers mean and how they are used.
- **Computation (B):** Students understand and demonstrate computation skills.
- **Data analysis and statistics (C):** Students understand and apply concepts of data analysis.
- **Probability (D):** Students understand and apply concepts of probability.
- **Geometry (E):** Students understand and apply concepts from geometry.
- **Measurement (F):** Students understand and demonstrate measurement skills.
- **Patterns, relations, and functions (G):** Students understand that mathematics is the science of patterns, relationships, and functions.
- **Algebra concepts (H):** Students understand and apply algebraic concepts.
- **Discrete mathematics (I):** Students understand and apply concepts in discrete mathematics.
- **Mathematical reasoning:** Students understand and apply concepts of mathematical reasoning.
- **Mathematical communication:** Students reflect upon and clarify their understanding of mathematical ideas and relationships.

These standards were used to create a reporting category framework for mathematics, shown below. The framework was divided into two major areas:

- **content**, which refers to the student's knowledge and conceptual and procedural understanding of each standard, and
- **application**, which refers to a student's use of knowledge and conceptual and procedural understanding as a basis for application through reasoning, inquiry, communication of ideas, and problem solving.

Each question in the mathematics assessment measured a content standard; in addition, each question was reported as measuring either content or application.

As shown in the table below, the goal for distribution of questions, or emphasis, across standards varied from grade to grade.

Content Standard	Grade		
	4	8	11
A. Number and Number Sense	15%	14%	10%
B. Computation	15%	11%	5%
C. Data Analysis and Statistics	12%	11%	10%
D. Probability	8%	11%	10%
E. Geometry	12%	11%	15%
F. Measurement	12%	10%	10%
G. Patterns, Relations, Functions	12%	13%	15%
H. Algebra Concepts	9%	14%	15%
I. Discrete Mathematics	5%	5%	10%

CONTENT SPECS

For students to function effectively as mathematical problem solvers, they must be taught how to apply and communicate basic concepts and procedures as well as how to do the procedures themselves.

Content questions measure what students have been taught directly. Included in these are the basic concepts and procedural skills from all the content standards. For example, in the numbers and number sense standard and the

computation standard, conceptual and procedural knowledge includes understanding of place value in our number system; the computational algorithms as applied to whole numbers, fractions, and decimals; and the concepts of ratio, proportion, and percent. In the data analysis and statistics standard, conceptual and procedural knowledge includes the reading of charts and graphs as well as the concepts of averages (means, medians, and modes) and the methods for computing them. Contextual settings used in questions measuring this category were very simple and were directly related to those used in the teaching of the concepts and the procedures.

Application questions measure what the students can do with what they have been taught. Included are questions requiring students to combine the basic concepts and procedures to solve real-life and mathematical problems, to evaluate their own ideas and the ideas of others using mathematical reasoning, and to communicate their ideas using the wealth of symbolic, pictorial, graphic, and verbal representations available in mathematics.

It is important to understand that application questions also measure mastery of the basic concepts and procedures. For example, in mathematics, 20 percent of the questions were either constructed- or extended-response questions (see "Item Types" on the next page), which were worth up to four and eight score points, respectively. In most cases, portions of these questions required the student to perform some problem solving, reasoning, and/or communicating, and so the questions were classified under applications. At the same time, however, the questions required the students to demonstrate their understanding of mathematics content. If a student did not show mastery of all aspects of a constructed- or extended-response question, or if he/she made careless errors, the student did not earn the highest score for that question. Thus, it can be said that **all** mathematics questions in the MEA measured content; some questions went beyond that realm, however, and were classified for reporting purposes as application.

ITEM TYPES

The MEA mathematics assessment included multiple-choice, short-answer, constructed-response, and extended-response questions. Short-answer questions, which were new in the revised MEA, required students to perform a computation or solve a simple problem. Extended-response questions in mathematics are similar to constructed-response questions except that they are more complex, requiring ten to twenty minutes of response time. Each type of question was worth a specific number of points in the student's total mathematics score, as shown below.

Type of Question	Possible Score Points
Multiple-Choice	0–1
Short-Answer	0–2
Constructed-Response	0–4
Extended-Response	0–8

TEST DESIGN

The tables below summarize the numbers and types of questions that were used in the MEA mathematics assessment for 1999–00. The tables show the construction of the common, matrix-sampled, and pretest portions of the assessment.

GRADE 4

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
4A (NC)	4	3	2		2	1			30
4B (C)	8	1	1		2		1		30
4C (C)	8	1	2		1	1			30

GRADES 8/11

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
4A NC	15	4	2		2	1			50
4B (C)	5	1	1	1	3	1	1		50

Key

- (C) = calculator use allowed
- (NC) = no calculator use allowed
- MC = multiple-choice questions
- SA = short-answer questions
- CR = constructed-response questions
- ER = extended-response questions

THE USE OF CALCULATORS IN THE MEA

The Maine educators who designed and developed the assessment test acknowledge the importance of mastering arithmetic algorithms. At the same time, they understand that the use of calculators is a necessary and important skill in society today. Calculators can save time and prevent error in the measurement of some higher-order thinking skills and allow students to do more sophisticated and intricate problems. For these reasons, it was decided that calculators should be permitted on some parts of the MEA mathematics assessment and prohibited on others. (Students were allowed to use any calculator with which they were familiar.)

The charts on the following pages outline the total number of possible points—as reported—by learning results and item type.

MATHEMATICS
Number of Points Possible
Grade 4

Standard	Common					Matrix Per Form					Total Possible Points
	MC	SA	CR	Points	Percent	MC	SA	CR	Points	Percent	
Content	18	4	0	22	25	49	18	0	67	75	89
Application	2	6	16	24	21	10	30	48	88	79	112
Numbers and Number Sense (Standard A)	4	0	4	8	27	14	4	4	22	73	30
Computation (Standard B)	1	4	0	5	17	7	14	4	25	83	30
Data Analysis and Statistics (Standard C)	1	2	4	7	33	6	0	8	14	67	21
Probability (Standard D)	4	0	0	4	29	2	0	8	10	71	14
Geometry (Standard E)	2	0	0	2	12.5	8	2	4	14	87.5	16
Measurement (Standard F)	1	2	4	7	23	11	8	4	23	77	30
Patterns, Relations, Functions (Standard G)	3	0	4	7	29	5	8	4	17	71	24
Algebra Concepts (Standard H)	2	2	0	4	19	3	6	8	17	81	21
Discrete Mathematics (Standard I)	2	0	0	2	13	3	6	4	13	87	15

MATHEMATICS
Number of Points Possible
Grade 8

Standard	Common						Matrix Per Form						Total Possible Points
	MC	SA	CR	ER	Points	Percent	MC	SA	CR	ER	Points	Percent	
Content	17	8	4	0	29	25	46	22	20	0	88	75	117
Application	3	2	8	8	21	24	14	26	28	0	68	76	89
Numbers and Number Sense (Standard A)	0	0	0	8	8	29	10	6	4	0	20	71	28
Computation (Standard B)	3	2	0	0	5	22	8	6	4	0	18	78	23
Data Analysis and Statistics (Standard C)	3	2	0	0	5	23	7	6	4	0	17	77	22
Probability (Standard D)	1	0	4	0	5	22	6	4	8	0	18	78	23
Geometry (Standard E)	3	2	0	0	5	21	3	4	12	0	19	79	24
Measurement (Standard F)	5	0	0	0	5	23	3	6	8	0	17	77	22
Patterns, Relations, Functions (Standard G)	2	0	4	0	6	23	8	8	4	0	20	77	26
Algebra Concepts (Standard H)	0	4	0	4	8	28	11	6	4	0	21	72	29
Discrete Mathematics (Standard I)	3	0	0	0	3	33	4	2	0	0	6	67	9

MATHEMATICS
Number of Points Possible
Grade 11

Standard	Common						Matrix Per Form						Total Possible Points
	MC	SA	CR	ER	Points	Percent	MC	SA	CR	ER	Points	Percent	
Content	14	4	0	0	18	27	30	6	12	0	48	73	66
Application	6	6	12	8	32	23	30	42	36	0	108	77	140
Numbers and Number Sense (Standard A)	4	0	0	0	4	24	9	4	0	0	13	76	17
Computation (Standard B)	1	2	0	0	3	20	6	2	4	0	12	80	15
Data Analysis and Statistics (Standard C)	1	0	4	0	5	22	8	6	4	0	18	78	23
Probability (Standard D)	3	2	0	0	5	23	7	2	8	0	17	77	22
Geometry (Standard E)	2	0	0	8	10	26	8	8	12	0	28	74	38
Measurement (Standard F)	1	2	0	0	3	13	6	6	8	0	20	87	23
Patterns, Relations, Functions (Standard G)	4	0	4	0	8	25	8	8	8	0	24	75	32
Algebra Concepts (Standard H)	3	0	4	0	7	26	8	8	4	0	20	74	27
Discrete Mathematics (Standard I)	1	4	0	0	5	56	0	4	0	0	4	44	9

CHAPTER 6—DESIGN OF THE SCIENCE AND TECHNOLOGY ASSESSMENT

BLUEPRINT

The science and technology framework was based on Maine's *Learning Results*, which identifies thirteen **content standards**, as listed below:

- **Classifying life forms (A):** Students understand that there are similarities within the diversity of all living things.
- **Ecology (B):** Students understand how living things depend on one another and on non-living aspects of the environment.
- **Cells (C):** Students understand that cells are the basic units of life.
- **Continuity and change (D):** Students understand the basis for all life and that all living things change over time.
- **Structure of matter (E):** Students understand the structure of matter and the changes it can undergo.
- **The Earth (F):** Students gain knowledge about the Earth and the processes that change it.
- **The universe (G):** Students gain knowledge about the universe and how humans have learned about it, and the principles upon which it operates.
- **Energy (H):** Students understand concepts of energy.
- **Motion (I):** Students understand the motion of objects and how forces can change that motion.
- **Inquiry and problem solving (J):** Students apply inquiry and problem-solving approaches in science and technology.
- **Scientific reasoning (K):** Students learn to formulate and justify ideas and to make informed decisions.
- **Communication (L):** Students communicate effectively in the applications of science and technology.
- **Implications of science and technology (M):** Students understand the historical, social, economic, environmental, and ethical implications of science and technology.

Nine of these standards (A through I) address the various content areas in science and technology; the remaining four (J, K, L, and M) highlight scientific applications. These were adapted and combined to create the reporting category framework for science and technology, shown below.

Content Standard	Application			
	J. Inquiry and Problem Solving	K. Scientific Reasoning	L. Communication	M. Implications of Science & Technology
A. Classifying Life Forms				
B. Ecology				
C. Cells				
D. Continuity and Change				
E. Structure of Matter				
F. The Earth				
G. The Universe				
H. Energy				
I. Motion				

All questions in the science and technology assessment measured a content standard; approximately 40 percent of the questions were written to measure a performance indicator in applications.

APPLICATIONS

The score for applications refers to a student's use of knowledge and conceptual and procedural understandings as a basis for application through reasoning, inquiry, communication of ideas, and problem solving.

ITEM TYPES

The MEA science and technology assessment included multiple-choice, short-answer, constructed-response, and extended-response questions. Short-answer questions, which were new in the revised MEA, required students to formulate an answer using one or two words or a short phrase. Extended-response questions in science and technology are similar to constructed-response questions except that they are more complex, requiring ten to twenty minutes of response time. Each type of question was worth a specific number of points in the student's total science and technology score, as shown on the next page.

Type of Question	Possible Score Points
Multiple-Choice	0–1
Short-Answer	0–2
Constructed-Response	0–4
Extended-Response	0–8

The scoring of extended response questions utilized either two four-point guides, one measuring science content and one measuring science applications, or one eight-point guide, measuring solely content or applications.

TEST DESIGN

The tables below summarize the numbers and types of questions that were used in the MEA science and technology assessment for 1999–00.

GRADE 4

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
2A	7	1	2		2	1			30
2B	7	2	1		2		1		30
2C	6	2	2		2				30

GRADES 8/11

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
2A	13	1	1	1	2	1			50
2B	7	4	2		4		1		50

Key

- MC = multiple-choice questions
- SA = short-answer questions
- CR = constructed-response questions
- ER = extended-response questions

The charts on the following pages outline the total number of possible points—as reported—by learning results and item type.

SCIENCE AND TECHNOLOGY
Number of Points Possible
Grade 4

Standard	Common					Matrix					Total Possible Points
	MC	SA	CR	Points	Percent	MC	SA	CR	Points	Percent	
Content	12	6	12	30	26	43	14	28	85	74	115
Classifying Life Forms (Standard A)	1	0	0	1	9	4	2	4	10	91	11
Ecology (Standard B)	2	0	0	2	17	6	0	4	10	83	12
Cells (Standard C)	2	2	0	4	27	5	2	4	11	73	15
Continuity and Change (Standard D)	1	0	4	5	42	5	2	0	7	58	12
Structure of Matter (Standard E)	1	0	0	1	10	3	2	4	9	90	10
The Earth (Standard F)	1	0	4	5	45	4	2	0	6	55	11
The Universe (Standard G)	2	2	0	4	24	7	2	4	13	76	17
Energy (Standard H)	1	0	0	1	8	5	2	4	11	92	12
Motion (Standard I)	1	2	4	7	47	4	0	4	8	53	15
Application	8	4	8	20	25	29	10	20	59	75	79
Inquiry and Problem Solving (Standard J)	4	2	8	14	61	7	2	0	9	39	23
Scientific Reasoning (Standard K)	1	0	0	1	5	7	8	4	19	95	20
Communication (Standard L)	2	2	0	4	15	10	0	12	22	85	26
Implications of Science and Technology (Standard M)	1	0	0	1	10	5	0	4	9	90	10

SCIENCE AND TECHNOLOGY
Number of Points Possible
Grade 8

Standard	Common						Matrix						Total Possible Points
	MC	SA	CR	ER	Points	Percent	MC	SA	CR	ER	Points	Percent	
Content	18	4	4	4	30	22	64	22	20	0	106	78	136
Classifying Life Forms (Standard A)	2	0	0	0	2	18	7	2	0	0	9	82	11
Ecology (Standard B)	1	0	0	4	5	36	7	2	0	0	9	64	14
Cells (Standard C)	3	0	0	0	3	17	7	4	4	0	15	83	18
Continuity and Change (Standard D)	2	0	0	0	2	12	9	2	4	0	15	88	17
Structure of Matter (Standard E)	2	2	0	0	4	22	6	4	4	0	14	78	18
The Earth (Standard F)	2	2	4	0	8	40	6	2	4	0	12	60	20
The Universe (Standard G)	2	0	0	0	2	13	7	2	4	0	13	87	15
Energy (Standard H)	2	0	0	0	2	20	6	2	0	0	8	80	10
Motion (Standard I)	2	0	0	0	2	15	9	2	0	0	11	85	13
Application	2	6	8	4	20	34	8	2	28	0	38	66	58
Inquiry and Problem Solving (Standard J)	2	0	0	0	2	11	5	0	12	0	17	89	19
Scientific Reasoning (Standard K)	0	2	0	4	6	33	2	2	8	0	12	67	18
Communication (Standard L)	0	2	4	0	6	55	1	0	4	0	5	45	11
Implications of Science and Technology (Standard M)	0	2	4	0	6	60	0	0	4	0	4	40	10

SCIENCE AND TECHNOLOGY
Number of Points Possible
Grade 11

Standard	Common						Matrix						Total Possible Points
	MC	SA	CR	ER	Points	Percent	MC	SA	CR	ER	Points	Percent	
Content	14	8	4	4	30	22	62	22	20	0	104	78	134
Classifying Life Forms (Standard A)	1	0	0	0	1	12.5	5	2	0	0	7	87.5	8
Ecology (Standard B)	1	0	0	4	5	36	7	2	0	0	9	64	14
Cells (Standard C)	1	0	0	0	1	6	7	4	4	0	15	94	16
Continuity and Change (Standard D)	3	2	0	0	5	36	7	2	0	0	9	64	14
Structure of Matter (Standard E)	1	2	0	0	3	18	8	2	4	0	14	82	17
The Earth (Standard F)	2	0	4	0	6	26	9	4	4	0	17	74	23
The Universe (Standard G)	2	2	0	0	4	27	5	2	4	0	11	73	15
Energy (Standard H)	1	0	0	0	1	10	7	2	0	0	9	90	10
Motion (Standard I)	2	2	0	0	4	24	7	2	4	0	13	76	17
Application	6	2	8	4	20	33	10	2	28	0	40	67	60
Inquiry and Problem Solving (Standard J)	2	0	4	0	6	43	4	0	4	0	8	57	14
Scientific Reasoning (Standard K)	0	0	0	0	0	0	0	2	8	0	10	100	10
Communication (Standard L)	3	0	0	4	7	35	5	0	8	0	13	65	20
Implications of Science and Technology (Standard M)	1	2	4	0	7	44	1	0	8	0	9	56	16

CHAPTER 7—DESIGN OF THE SOCIAL STUDIES ASSESSMENT

BLUEPRINT

The social studies framework was based on Maine's *Learning Results*, which identifies a total of thirteen **content standards** in the four disciplines—civics and government; history; geography; and economics—as listed below:

CIVICS AND GOVERNMENT

- **Rights, responsibilities, and participation:** Students understand the rights and responsibilities of civic life and employ the skills of effective civic participation.
- **Purpose and types of government:** Students understand the types and purposes of governments, their evolution, and their relationships with the governed.
- **Fundamental principles of government and constitutions:** Students understand the constitutional principles and the democratic foundations of the political institutions of the United States.
- **International relations:** Students understand the political relationships among the United States and other nations.

HISTORY

- **Chronology:** Students use the chronology of history and major eras to demonstrate the relationships of events and people.
- **Historical knowledge, concepts, and patterns:** Students develop historical knowledge of major events, people, and enduring themes in the United States, in Maine, and throughout world history.
- **Historical inquiry, analysis, and interpretation:** Students learn to evaluate resource material such as documents, artifacts, maps, artwork, and literature and to make judgments about the perspectives of the authors and their credibility when interpreting current historical events.

GEOGRAPHY

- **Skills and tools:** Students know how to construct and interpret maps and use globes and other geographic tools to locate and derive information about people, places, regions, and environments.

- **Human interaction with environments:** Students understand and analyze the relationships among people and their physical environments.

ECONOMICS

- **Personal and consumer economics:** Students understand that economic decisions are based on the availability of resources and the costs and benefits of choices.
- **Economic systems of the United States:** Students understand the economic system of the United States, including its principles, development, and institutions.
- **Comparative systems:** Students analyze how different economic systems function and change over time.
- **International trade and global interdependence:** Students understand the patterns and results of international trade.

These thirteen standards were used to create the reporting category framework for social studies, shown below.

Standard	Percentage of Questions Emphasizing Content	Percentage of Questions Emphasizing Application
Civics and Government:		
A. Rights, Responsibilities, and Participation	50%	50%
B./C. Purposes, Types, and Fundamental Principles	60%	40%
D. International Relations	60%	40%
History:		
A./B. Chronology and Historical Knowledge, Concepts, and Patterns	60%	40%
C. Historical Inquiry, Analysis, and Interpretation	40%	60%
Geography:		
A. Skills and Tools	40%	60%
B. Human Interaction with Environments	60%	40%
Economics:		
A. Personal and Consumer Economics	50%	50%
B./C. Economic Systems	50%	50%
D. International Trade and Global Interdependence (Grades 8 and 11)	60%	40%

Social studies education stresses a strong commitment to content knowledge, emphasizes the student's ability to engage in complex thinking and reasoning skills, and emphasizes the clear communication of ideas. Social studies assessment focuses on both content and applications to evaluate what students know and can demonstrate.

ITEM TYPES

The MEA social studies assessment included multiple-choice, short-answer, constructed-response, and extended-response questions. Short-answer questions, which were new in the revised MEA, required students to answer questions using one or two words or a short phrase. Extended-response questions in social studies are similar to constructed-response questions except that they are more complex, requiring ten to twenty minutes of response time. Each type of question was worth a specific number of points in the student's total social studies score, as shown below.

Type of Question	Possible Score Points
Multiple-Choice	0–1
Short-Answer	0–2
Constructed-Response	0–4
Extended-Response	0–8

TEST DESIGN

The tables below summarize the numbers and types of common, matrix-sampled, and pretest questions that were used in the 1999–00 social studies assessment.

GRADE 4

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
3A	7	1	2		2	1			30
3B	7	2	1		2		1		30
3C	6	2	2		2				30

GRADES 8/11

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
3A	13	1	1	1	2	1			50
3B	7	4	2		4		1		50

Key

- MC = multiple-choice questions
- SA = short-answer questions
- CR = constructed-response questions
- ER = extended-response questions

The charts on the following pages outline the total number of possible points—as reported—by learning results and item type.

SOCIAL STUDIES
Number of Points Possible
Grade 4

Standard	Common					Matrix					Total Possible Points
	MC	SA	CR	Points	Percent	MC	SA	CR	Points	Percent	
Content	17	4	0	21	20	69	6	8	83	80	104
Application	3	6	20	29	32	3	18	40	61	68	90
Civics and Government (Standards A and B)	6	2	4	12	21	26	6	12	44	79	56
Rights, Responsibilities, and Participation (Standard A)	4	2	4	10	37	11	2	4	17	63	27
Purpose, Types, and Fundamental Principles of Government (Standard B)	1	0	0	1	5	8	4	8	20	95	21
History (Standards A, B, and C)	5	2	4	11	19	24	6	16	46	81	57
Chronology, Historical Knowledge, Concepts, and Patterns (Standards A and B)	4	2	4	10	26	16	4	8	28	74	38
Historical Inquiry, Analysis, and Interpretation (Standard C)	1	0	0	1	7	8	2	4	14	93	15
Geography (Standards A and B)	4	4	8	16	28	17	8	16	41	72	57
Skills and Tools (Standard A)	2	0	0	2	8	12	2	8	22	92	24
Human Interaction with Environments (Standard B)	2	4	8	14	42	5	6	8	19	58	33
Economics (Standards A and B)	5	2	4	11	46	5	4	4	13	54	24
Personal and Consumer Economics/ Economic Systems (Standards A and B)	5	2	4	11	100	0	0	0	0	0	11

SOCIAL STUDIES
Number of Points Possible
Grade 8

Standard	Common						Matrix						Total Possible Points
	MC	SA	CR	ER	Points	Percent	MC	SA	CR	ER	Points	Percent	
Content	20	8	4	0	32	26	72	18	0	0	90	74	122
Application	0	2	8	8	18	25	0	6	48	0	54	75	72
Civics and Government (Standards A, B, C, and D)	5	4	4	0	13	27	18	6	12	0	36	73	49
Rights, Responsibilities, and Participation (Standard A)	2	0	4	0	6	37.5	6	0	4	0	10	62.5	16
Purpose, Types, and Fundamental Principles of Government (Standards B and C)	2	2	0	0	4	20	8	4	4	0	16	80	20
International Relations (Standard D)	1	2	0	0	3	23	4	2	4	0	10	77	13
History (Standards A, B, and C)	5	2	0	0	7	14	21	6	16	0	43	86	50
Chronology, Historical Knowledge, Concepts, and Patterns (Standards A and B)	5	2	0	8	15	36	19	4	4	0	27	64	42
Historical Inquiry, Analysis, and Interpretation (Standard C)	0	0	0	0	0	0	2	2	8	0	12	100	12
Geography (Standards A and B)	6	2	4	8	20	36	18	6	12	0	36	64	56
Skills and Tools (Standard A)	3	2	0	0	5	36	7	2	0	0	9	64	14
Human Interaction with Environments (Standard B)	3	0	4	0	7	21	11	4	12	0	27	79	34
Economics (Standards A, B, and D)	4	2	4	0	10	26	15	6	8	0	29	74	39
Personal and Consumer Economics (Standard A)	2	0	0	0	2	17	4	2	4	0	10	83	12
Economic Systems/Comparative Systems (Standards B and C)	0	2	4	0	6	32	9	4	0	0	13	68	19
International Trade and Global Interdependence (Standard D)	2	0	0	0	2	25	2	0	4	0	6	75	8

SOCIAL STUDIES
Number of Points Possible
Grade 11

Standard	Common						Matrix						Total Possible Points
	MC	SA	CR	ER	Points	Percent	MC	SA	CR	ER	Points	Percent	
Content	15	10	0	0	25	25	53	18	4	0	75	75	100
Application	5	0	12	8	25	27	19	6	44	0	69	73	94
Civics and Government (Standards A, B, C, and D)	6	2	4	0	12	25	16	8	12	0	36	75	48
Rights, Responsibilities, and Participation (Standard A)	3	0	0	0	3	27	2	2	4	0	8	73	11
Purpose, Types, and Fundamental Principles of Government (Standards B and C)	3	0	4	0	7	24	10	4	8	0	22	76	29
International Relations (Standard D)	0	2	0	0	2	25	4	2	0	0	6	75	8
History (Standards A, B, and C)	5	4	0	8	17	28	23	4	16	0	43	72	60
Chronology, Historical Knowledge, Concepts, and Patterns (Standards A and B)	5	2	0	8	15	30	23	4	8	0	35	70	50
Historical Inquiry, Analysis, and Interpretation (Standard C)	0	2	0	0	2	20	0	0	8	0	8	80	10
Geography (Standards A and B)	5	2	4	0	11	23	18	6	12	0	36	77	47
Skills and Tools (Standard A)	3	0	0	0	3	15	7	2	8	0	17	85	20
Human Interaction with Environments (Standard B)	2	2	4	0	8	30	11	4	4	0	19	70	27
Economics (Standards A, B, and D)	4	2	4	0	10	26	15	6	8	0	29	74	39
Personal and Consumer Economics (Standards A)	1	2	0	0	3	27	2	2	4	0	8	73	11
Economic Systems/Comparative Systems (Standards B and C)	2	0	4	0	6	33	10	2	0	0	12	67	18
International Trade and Global Interdependence (Standard D)	1	0	0	0	1	10	3	2	4	0	9	90	10

CHAPTER 8—DESIGN OF THE VISUAL AND PERFORMING ARTS ASSESSMENT

BLUEPRINT

The visual and performing arts assessment included four disciplines: dance, music, theater, and visual arts. The arts framework was based on Maine's *Learning Results*, which identifies three **content standards** in the arts, as listed below:

- **Creative expression (A):** Students create and/or perform to express ideas and feelings.
- **Cultural heritage (B):** Students understand the cultural contributions (social, ethical, political, and religious dimensions) of the arts, understand how the arts shape and are shaped by prevailing cultural and social beliefs and values, and recognize exemplary works from a variety of cultures and historical periods.
- **Criticism and aesthetics (C):** Students reflect upon and assess the characteristics and merits of works of art.

These three standards were used to create the reporting category framework for the visual and performing arts, shown below.

Discipline	Standard		
	A. Creative Expression	B. Cultural Heritage	C. Criticism and Aesthetics
Dance			
Music			
Theater			
Visual Arts			

Each row and each column of the framework constitutes a reporting category for school- and district-level results in the MEA—for example, music/cultural heritage. Student-level results were not reported in the visual and performing arts, as no common items were used in this area.

It should be noted that not all of the performance indicators associated with each content standard (see *Learning Results*) can be assessed reliably and validly using a paper-and-pencil test. For example, some of the performance indicators included under the standard for creative expression would best be measured in other ways. For this reason, additional methods of assessment for these performance indicators are being studied.

The distribution of questions, or emphasis, across the arts disciplines in the MEA varied from one grade level to another, as shown in the table below.

Discipline	Grade		
	4	8	11
Danse	10%	10%	15%
Music	40%	40%	35%
Theater	10%	10%	15%
Visual Arts	40%	40%	35%

ITEM TYPES

The MEA visual and performing arts assessment included multiple-choice, and constructed-response questions. Each type of question was worth a specific number of points, as shown below.

Type of Question	Possible Score Points
Multiple Choice	0–1
Constructed Response	0–4

TEST DESIGN

The table below summarizes the numbers and types of matrix-sampled questions that were used in the 1999-00 visual and performing arts assessment.

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
5A					6		1		30

Key

- MC = multiple-choice questions
- SA = short-answer questions
- CR = constructed-response questions
- ER = extended-response questions

The charts on the following pages outline the total number of possible points—as reported—by learning results and item type.

VISUAL AND PERFORMING ARTS
Number of Points Possible
Grade 4

Standard	Common				Matrix					Total Possible Points
	MC	SA	CR	Points	MC	SA	CR	Points	Percent	
Dance					14	0	12	26	100	26
Music					22	0	12	34	100	34
Theater					14	0	12	26	100	26
Visual Arts					22	0	12	34	100	34
Creative Expression (Standard A)					33	0	16	49	100	49
Cultural Heritage (Standard B)					19	0	8	27	100	27
Criticism and Aesthetics (Standard C)					20	0	24	44	100	44

VISUAL AND PERFORMING ARTS
Number of Points Possible
Grade 8

Standard	Common				Matrix					Total Possible Points
	MC	SA	CR	Points	MC	SA	CR	Points	Percent	
Dance					10	0	12	22	100	22
Music					25	0	12	37	100	37
Theater					12	0	12	24	100	24
Visual Arts					25	0	12	37	100	37
Creative Expression (Standard A)					28	0	16	44	100	44
Cultural Heritage (Standard B)					23	0	20	43	100	43
Criticism and Aesthetics (Standard C)					21	0	12	33	100	33

VISUAL AND PERFORMING ARTS
Number of Points Possible
Grade 11

Standard	Common				Matrix					Total Possible Points
	MC	SA	CR	Points	MC	SA	CR	Points	Percent	
Dance					11	0	12	23	100	23
Music					25	0	12	37	100	37
Theater					12	0	12	24	100	24
Visual Arts					24	0	12	36	100	36
Creative Expression (Standard A)					27	0	20	47	100	47
Cultural Heritage (Standard B)					21	0	8	29	100	29
Criticism and Aesthetics (Standard C)					24	0	20	44	100	44

CHAPTER 9—DESIGN OF THE HEALTH EDUCATION ASSESSMENT

BLUEPRINT

The health framework was based on Maine's *Learning Results*, which identifies six **content standards**, as shown below:

- **Health concepts:** Students understand health promotion and disease prevention concepts.
- **Health information, services, and products:** Students know how to acquire valid information about health issues, services, and products.
- **Health promotion and risk reduction:** Students understand how to reduce their health risks through the practice of healthy behaviors.
- **Influences on health:** Students understand how media techniques, cultural perspectives, technology, peers, and family influence behaviors that affect health.
- **Communication skills:** Students understand that skillful communication can contribute to better health for them, their families, and the community.
- **Decision making and goal setting:** Students learn how to set personal goals and make decisions that lead to better health.

These six standards were combined with the ten health education content areas identified by the 1984 Education Reform Act to create a reporting category framework for health, as shown on the next page.

Content Area	Health Standard					
	A. Health Concepts	B. Health Information, Services, and Products	C. Health Promotion and Risk Reduction	D. Influences on Health	E. Communication Skills	F. Decision Making and Goal Setting
Community, Consumer, and Environmental Health						
Personal and Nutritional Health						
Family Life Education and Growth and Development						
Safety and Injury Prevention						
Tobacco, Alcohol, and Other Drug Use Prevention						
Prevention and Control of Disease and Disorders						
Total	30%	70%				

Thirty percent of the questions measured health standard (A); they were divided among the six content areas. The remaining 70 percent of the questions were divided among the other five health standards (B through F) and the six content areas. The distribution of questions was 10 percent to 20 percent for each standard, determined by its developmental appropriateness for the specific grade being assessed.

A portion of the questions in the health assessment were developed by the Health Education Assessment Project for the State Collaborative on Assessment and Student Standards (SCASS) under the auspices of the Council of Chief State School Officers. Each SCASS question that was used or adapted was aligned with a performance indicator from Maine's health education standards. Maine educators on the Development Advisory Committee developed the remainder of the questions.

ITEM TYPES

The MEA health assessment included multiple-choice, short-answer, constructed-response, and extended-response questions. Short-answer questions, which were new in the revised MEA, required students to formulate answers using one or two words or a short phrase. Extended-response questions in health are similar to constructed-response

questions except that they are more complex, requiring ten to twenty minutes of response time. Each type of question was worth a specific number of points in the student's total health score, as shown below.

Type of Question	Possible Score Points
Multiple-Choice	0–1
Short-Answer	0–2
Constructed-Response	0–4
Extended-Response	0–8

TEST DESIGN

At every grade level, the assessment included no common questions but was constructed solely of matrix-sampled questions. The tables below summarize the numbers and types of questions that were used in the 1999–00 health education assessment.

GRADE 4

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
4A					6	1	3		30

GRADES 8/11

Session	COMMON				MATRIX				Time (minutes)
	MC	SA	CR	ER	MC	SA	CR	ER	
4A					6	1	1	1	40

Key

- MC = multiple-choice questions
- SA = short-answer questions
- CR = constructed-response questions
- ER = extended-response questions

The charts on the following pages outline the total number of possible points—as reported—by learning results and item type.

HEALTH EDUCATION
Number of Points Possible
Grade 4

Standard	Common						Matrix						Total Possible Points
	MC	SA	CR	ER	Points	%	MC	SA	CR	ER	Points	%	
Health Concepts (Standard A)							15	2	24	0	41	100	41
Health Information, Services, and Products (Standard B)							16	6	36	0	58	100	58
Health Promotion and Risk Reduction (Standard C)							9	4	24	0	37	100	37
Influences on Health (Standard D)							9	6	28	0	43	100	43
Communication Skills (Standard E)							14	4	20	0	38	100	38
Decision Making and Goal Setting (Standard F)							9	2	12	0	23	100	23
Community, Consumer, and Environmental Health							21	8	44	0	73	100	73
Personal and Nutritional Health							13	4	8	0	25	100	25
Family Life Education and Growth and Development							22	8	20	0	50	100	50
Safety and Injury Prevention							8	0	16	0	24	100	24
Tobacco, Alcohol, and Other Drug Use Prevention							5	2	28	0	35	100	35
Prevention and Control of Disease and Disorders							3	2	28	0	33	100	33

HEALTH EDUCATION
Number of Points Possible
Grade 8

Standard	Common						Matrix						Total Possible Points
	MC	SA	CR	ER	Points	%	MC	SA	CR	ER	Points	%	
Health Concepts (Standard A)							8	4	16	0	28	100	28
Health Information, Services, and Products (Standard B)							17	4	4	32	57	100	57
Health Promotion and Risk Reduction (Standard C)							13	6	4	24	47	100	47
Influences on Health (Standard D)							12	2	12	16	42	100	42
Communication Skills (Standard E)							13	4	8	16	41	100	41
Decision Making and Goal Setting (Standard F)							9	4	4	8	25	100	25
Community, Consumer, and Environmental Health							47	10	0	24	81	100	81
Personal and Nutritional Health							7	0	12	8	27	100	27
Family Life Education and Growth and Development							10	6	4	16	36	100	36
Safety and Injury Prevention							1	6	16	8	31	100	31
Tobacco, Alcohol, and Other Drug Use Prevention							3	2	12	24	41	100	41
Prevention and Control of Disease and Disorders							3	0	4	16	23	100	23

HEALTH EDUCATION
Number of Points Possible
Grade 11

Standard	Common						Matrix						Total Possible Points
	MC	SA	CR	ER	Points	%	MC	SA	CR	ER	Points	%	
Health Concepts (Standard A)							10	4	8	16	38	100	38
Health Information, Services, and Products (Standard B)							12	4	8	12	36	100	36
Health Promotion and Risk Reduction (Standard C)							10	4	8	16	38	100	38
Influences on Health (Standard D)							15	4	8	8	35	100	35
Communication Skills (Standard E)							9	4	8	24	45	100	45
Decision Making and Goal Setting (Standard F)							15	4	8	16	43	100	43
Community, Consumer, and Environmental Health							31	6	8	36	81	100	81
Personal and Nutritional Health							7	10	4	4	25	100	25
Family Life Education and Growth and Development							12	2	8	12	34	100	34
Safety and Injury Prevention							7	6	4	24	41	100	41
Tobacco, Alcohol. and Other Drug Use Prevention							3	0	12	16	31	100	31
Prevention and Control of Disease and Disorders							11	0	12	0	23	100	23

SECTION II: TEST ADMINISTRATION

CHAPTER 10—TEST ADMINISTRATION

RESPONSIBILITY FOR ADMINISTRATION

As indicated in the *Coordinator's Manual*, principals and/or their designated MEA coordinators were responsible for the proper administration of the MEA. Manuals and certification forms were used to ensure the uniformity of administration procedures from school to school.

PROCEDURES

Each principal and/or the school's designated MEA coordinator was instructed to read the *Coordinator's Manual* prior to testing and to be familiar with the instructions given in the *Test Administrator's Manual*. The *Coordinator's Manual* provided each school with checklists to help it prepare for testing. The checklists outlined tasks for the schools to perform before, during, and after test administration. Along with these checklists, the *Coordinator's Manual* outlined the nature of the testing material being sent to each school, how to inventory the material, how to track it during administration, and how to return the material once testing was complete. It also contained information about including or excluding students. The *Test Administrator's Manual* included checklists for the administrators to prepare themselves, their classrooms, and their students for the administration of the test. The *Test Administrator's Manual* contained sections that detailed the procedure to be followed for each test session, and it contained instructions on preparing the material prior to giving it to the principal/coordinator for its return to Advanced Systems.

ADMINISTRATOR TRAINING

In addition to distributing the *Coordinator's* and *Test Administrator's Manual*s, the Maine Department of Education along with Advanced Systems conducted two ITV workshops (one in the fall and one in the winter) to train and inform school personnel about the new MEA.

STATE PARTICIPATION RATES—FALL 1999

GRADE 4

Students Excluded from Report(s):	Number of Students	Percentage of Students
students totally excluded from testing (took no session of the assessment) due to an identified disability	279	2%
students partially excluded from testing (excluded from some but not all sessions of the assessment) due to an identified disability	358	2%
students tested, but excluded from report because they receive special education and related services for more than 60% of the school day in a composite or self-contained program (categories 24 or 25 on EF-S-204)	73	0%
students totally excluded from testing because of LEP, Title 1 decision or other approved reason	81	0%
students partially excluded from testing because of LEP, Title 1 decision or other approved reason	19	0%
others totally excluded from testing	236	1%
others partially excluded from testing	351	2%
Students with Identified Disability Completing All Subjects without Accommodations	1390	8%
Students with Identified Disability Completing All Subjects with Accommodations	10	0%

GRADE 8

Students Excluded from Report(s):	Number of Students	Percentage of Students
students totally excluded from testing (took no session of the assessment) due to an identified disability	154	1%
students partially excluded from testing (excluded from some but not all sessions of the assessment) due to an identified disability	117	1%
students tested, but excluded from report because they receive special education and related services for more than 60% of the school day in a composite or self-contained program (categories 24 or 25 on EF-S-204)	196	1%
students totally excluded from testing because of LEP, Title 1 decision or other approved reason	73	0%
students partially excluded from testing because of LEP, Title 1 decision or other approved reason	12	0%
others totally excluded from testing	252	1%
others partially excluded from testing	486	3%
Students with Identified Disability Completing All Subjects without Accommodations	1607	9%
Students with Identified Disability Completing All Subjects with Accommodations	5	0%

GRADE 11

Students Excluded from Report(s):	**Number of Students**	**Percentage of Students**
students totally excluded from testing (took no session of the assessment) due to an identified disability	153	1%
students partially excluded from testing (excluded from some but not all sessions of the assessment) due to an identified disability	20	0%
students tested, but excluded from report because they receive special education and related services for more than 60% of the school day in a composite or self-contained program (categories 24 or 25 on EF-S-204)	62	0%
students totally excluded from testing because of LEP, Title 1 decision or other approved reason	190	1%
students partially excluded from testing because of LEP, Title 1 decision or other approved reason	8	0%
others totally excluded from testing	164	1%
others partially excluded from testing	701	5%
Students with Identified Disability Completing All Subjects without Accommodations	841	6%
Students with Identified Disability Completing All Subjects with Accommodations	10	0%

STATE PARTICIPATION RATES—SPRING 2000

GRADE 4

Students Excluded from Report(s):	**Number of Students**	**Percentage of Students**
students totally excluded from testing (took no session of the assessment) due to an identified disability	241	1%
students partially excluded from testing (excluded from some but not all sessions of the assessment) due to an identified disability	64	0%
students tested, but excluded from report because they receive special education and related services for more than 60% of the school day in a composite or self-contained program (categories 24 or 25 on EF-S-204)	118	1%
students totally excluded from testing because of LEP, Title 1 decision or other approved reason	46	0%
students partially excluded from testing because of LEP, Title 1 decision or other approved reason	1	0%
others totally excluded from testing	256	2%
others partially excluded from testing	223	1%
Students with Identified Disability Completing All Subjects without Accommodations	296	2%
Students with Identified Disability Completing All Subjects with Accommodations	1592	10%

GRADE 8

Students Excluded from Report(s):	Number of Students	Percentage of Students
students totally excluded from testing (took no session of the assessment) due to an identified disability	164	1%
students partially excluded from testing (excluded from some but not all sessions of the assessment) due to an identified disability	24	0%
students tested, but excluded from report because they receive special education and related services for more than 60% of the school day in a composite or self-contained program (categories 24 or 25 on EF-S-204)	233	1%
students totally excluded from testing because of LEP, Title 1 decision or other approved reason	54	0%
students partially excluded from testing because of LEP, Title 1 decision or other approved reason	4	0%
others totally excluded from testing	364	2%
others partially excluded from testing	335	2%
Students with Identified Disability Completing All Subjects without Accommodations	424	2%
Students with Identified Disability Completing All Subjects with Accommodations	1308	7%

GRADE 11

Students Excluded from Report(s):	Number of Students	Percentage of Students
students totally excluded from testing (took no session of the assessment) due to an identified disability	144	1%
students partially excluded from testing (excluded from some but not all sessions of the assessment) due to an identified disability	13	0%
students tested, but excluded from report because they receive special education and related services for more than 60% of the school day in a composite or self-contained program (categories 24 or 25 on EF-S-204)	78	1%
students totally excluded from testing because of LEP, Title 1 decision or other approved reason	173	1%
students partially excluded from testing because of LEP, Title 1 decision or other approved reason	7	0%
others totally excluded from testing	351	2%
others partially excluded from testing	396	3%
Students with Identified Disability Completing All Subjects without Accommodations	311	2%
Students with Identified Disability Completing All Subjects with Accommodations	565	4%

PARTICIPATION REQUIREMENTS

The following categories of students were allowed to be considered for modifications:

- Students who had an identified exceptionality/disability
- Students who had been identified as limited English proficient (LEP)
- Students who were unable to work independently in any of the subjects assessed
- Students who were ill or incapacitated in some way

All students who were considered for modifications on the MEA were to have had their individual situations reviewed by a group within the school prior to the time of testing. For every student with an identified exceptionality requiring an Individual Educational Plan (IEP), schools were required to hold a Pupil Evaluation Team (PET) meeting that addressed that student's needs for modifications. Other students needing test modifications who did not have an identified exceptionality were required to attend a meeting that included one of the student's teachers, the building principal, related services personnel, and, whenever possible, the student's parents. If it was not possible for the parents to attend the meeting, it was required that they be notified of the committee's recommendations for modifications prior to the time of testing.

Recommended modifications were to be consistent with those modifications already being employed in the student's instructional program. Any such modifications were reflected either in the minutes of the PET meeting (for students requiring an IEP) or in a statement prepared for the cumulative folders of students not requiring IEPs. The following is the suggested statement that schools were given as a model:

The student will/will not participate in the __th-grade Maine Educational Assessment as scheduled during the month of _____ 19__. The following test modifications will be observed: (list modifications)

EXCLUSION FROM THE ASSESSMENT

Exclusion was defined as the most extreme modification of the assessment. Since it was clear that the legislation's intent was to include as many students as possible, it was recommended that exclusion be considered only as a last resort.

On those occasions where it was deemed necessary to exclude a student from sections of the assessment or from the assessment as a whole, it was recommended that exclusion be limited to only those sections of the MEA that were considered inappropriate for that particular student. Exclusion was to be selected only after the various types of modifications available had been fully explored and it was felt that the assessment would not yield a valid indication of how a student functioned in a given content area. For example, even students who were reading two years below

grade level were advised to take the reading section because those scores would give a fair representation of their current level of functioning in reading. If, however, after examining all of the possible modifications, a local school decided that the assessment or sections of it would be inappropriate for a given student, that student could be excluded.

Students Enrolled in Ungraded or Multiage Programs

For the purposes of the assessment, it was recommended that students enrolled in ungraded or multiage programs be tested with the fourth grade if they were nine years old, with the eighth grade if they were thirteen, and with the eleventh grade if they were seventeen.

Documentation of Modifications or Exclusions

Information about the modifications given to students or the reasons for exclusion was to be provided on the front page of the student's response booklet. This information was to be coded in by staff, not students, after testing was completed. The *Coordinator's* and *Test Administrator's Manuals* provided directions on coding in the information related to modification(s), partial exclusion, and exclusion, and every student who was totally excluded had to be accounted for in the designated section of the response booklet.

Testing Irregularities

The following is a breakdown of the 1999–00 assessment irregularities:

GRADE	COMMON ITEMS	MATRIX ITEMS	TOTAL NUMBER OF IRREGULAR ITEMS
4	0	2	2
8	0	0	0
11	1*	0	1

*Form 6 only.

The charts on the following pages outline the irregular items on the test and an explanation of the errors.

GRADE 4 IRREGULARITIES

FORM	CONTENT	SESSION	Q- #	TYPE	Q-TYPE	DESCRIPTION OF ERROR
8	Mathematics	4B	25	M	CR	One of the shapes at the top of the page was an oval. In the box of clues, instead of an oval, a circle was used. Students were instructed that when answering this question, they should use the ovals and the circles as if they were the same shape.
10	Mathematics	4B	22	M	MC	There was no answer to this question. Students were instructed to leave the question blank and told that the question would not count.

GRADE 11 IRREGULARITIES

FORM	CONTENT	SESSION	Q- #	TYPE	Q-TYPE	DESCRIPTION OF ERROR
6	Mathematics	4A	19	M	SA	There should have been an arrow from Salt Lake City to Boston. Students were instructed to draw the arrow in their test booklets, starting with Salt Lake City and ending with Boston.

SECTION III: DEVELOPMENT AND REPORTING OF SCORES

CHAPTER 11—SCORING

MACHINE-SCORED ITEMS

Once the 1999–00 booklets had been logged in, identified with appropriate scannable, preprinted school information sheets, examined for extraneous materials, and batched, they were moved into the scanning area. For all response booklets (and questionnaires and other forms that required imaging/scanning) to be imaged, this area was the last stop in the processing loop in which the documents themselves were handled.

At that point, 100 percent of the response document and other scannable information necessary to produce the required reports had been captured and converted into an electronic format, including all student identification and demographics, selected-response answers, and digital image clips of handwritten responses. The digital image clip information allowed Advanced Systems to replicate student responses on the readers' monitors just as they had appeared on the originals. From that point on, the entire process—data processing, scoring, range-finding data analysis, and reporting—was accomplished without further reference to the originals.

The first step in that conversion was the removal of the booklet bindings so that the individual pages could pass through the scanners one at a time. Once cut, the sheets were put back in their proper boxes and placed in storage until needed for the scanning/imaging process.

Customized scanning programs for all scannables were prepared to selectively read the student response booklets and to format the scanned information electronically according to predetermined requirements. Any information (including multiple-choice response data) that had been designated time-critical or process-critical was handled first.

In addition to numerous real-time quality control checks, duplex read, a transport printer that prints a unique identifying number on each sheet of each booklet, and on-line editing capability, the new 5000i scanners offer features that make them compatible with Internet technology.

SCANNING QUALITY CONTROL

NCS scanners are equipped with many built-in safeguards that prevent data errors. The scanning hardware is continually monitored for conditions that will cause the machine to shut down if standards are not met. It will display an error message and prevent further scanning until the condition is corrected. The areas monitored include document page and integrity checks, user-designed on-line edits, and many internal checks of electronic functions.

Before every scanning shift begins, Advanced Systems operators perform a daily diagnostic routine. This is yet another step to protect data integrity and one that has been done faithfully for the many years that we have been involved in production scanning. In the rare event that the routine detects a photocell that appears to be out of range, we calibrate that machine and perform the test again. If the read is still not up to standard, we call for assistance from our field service engineer.

As a final safeguard, spot checks of scanned files, bubble by bubble and image by image, were routinely made throughout scanning runs. The result of these precautions, from the original layout of the scanning form to the daily vigilance of our operators, was a scan error rate well below .001.

ELECTRONIC DATA FILES

Once the data had been entered and the scanning logs and other paperwork completed, the booklets themselves were put into storage (where they stayed for at least 180 days beyond the close of the fiscal year). When it had been determined that the files were complete and accurate, those files were duplicated electronically and made available

for many other processing options. Completed files were loaded onto our local area network (LAN) for transfer to Advanced Systems' proprietary I-Score system for scoring. Those files were then used to identify (and print out) papers to be used in the range-finding and standard-setting processes, and the data was made transferable via the Internet, CD-ROM, or optical disk.

ITEMS SCORED BY READERS

Test and answer materials were handled as little as possible to minimize the possibility of loss, mishandling, or breach of security. Once scanned, either by optical mark reader or the I-Score system, papers were stored securely in areas with limited personnel access.

As explained in the following sections on scoring, the I-Score system itself ensures the security of responses and test items: all scoring is "blind"; that is, no student names are associated with viewed responses or raw scores and all scoring personnel are subject to the same nondisclosure requirements and supervision as regular Advanced Systems staff.

I-SCORE

After the 1999–00 test material had been loaded into the LAN, I-Score sent electronically scanned images of student work to individual readers at computer terminals who evaluated each response and recorded each student's score via keypad or mouse entry. When the reader had finished with one response, the next response appeared immediately on the computer screen. In that way, the system guaranteed complete anonymity of individual students and ensured the randomization of responses during scoring.

Although I-Score is based on conventional scoring techniques, it also offers numerous benefits, not the least of which is raising the bar on scoring process capability. Some of the benefits are as follows:

- real-time information on scorer reliability, read-behinds, and overall process monitoring;
- early access to subsets of data for tasks such as standard setting;
- reduced material handling, which not only saves time and labor, but also enhances the security of materials; and
- immediate access to samples of student responses and scores for reporting and analysis through electronic media.

Scoring operations, directed by the manager of scoring services, were carried out by a highly qualified staff. The staff included

- chief readers, who oversaw all training and scoring within particular subject areas;
- quality assurance coordinators (QACs), who led range-finding and training activities and monitored scoring consistency and rates;
- verifiers, who performed read-behinds of readers and assisted at scoring tables as necessary; and
- readers, who performed the bulk of the scoring.

Table 11-1 summarizes the qualifications of the 1999–00 MEA quality assurance coordinators and readers.

Table 11-1
Qualifications of 1999/00 MEA QACs and Readers

1999 Fall Administration					
Scoring Responsibility	Educational Credentials				Total
	Doctorate	Masters	Bachelors	Other	
QACs	10.53	68.42	21.05	0.00	100%
Readers	1.79	17.40	35.27	45.54	100%
2000 Spring Administration					
Scoring Responsibility	Educational Credentials				Total
	Doctorate	Masters	Bachelors	Other	
QACs	0.00	53.33	46.67	0.00	100%
Readers	4.89	14.66	39.85	40.60	100%

PRELIMINARY ACTIVITIES

Preliminary activities for scoring included (1) participating in the planning and design of documents to be used for scoring, (2) reviewing items and score guides for rangefinding and training and the creation of rangefinding packets, and (3) selecting scoring staff and training them for scoring.

Planning and Designing Documents

At the request of Advanced Systems' project manager, scoring personnel advised project management and DOE staff on the program design in order to support an efficient and effective scoring process. Scoring staff contributed also to the design of

- response documents and the image-capture process to yield acceptable image clips (also defining file format and layout); and
- scoring benchmarks composed of the guide, subject background information, and anchor papers.

Reviewing Items and Guides (Range finding)

Before the scheduled start of scoring activities, scoring center staff reviewed test items and scoring guides for range finding. At that point, chief readers and selected QACs prepared scorer training materials.

Advanced Systems' scoring staff (including test developers) selected one or two anchor examples for each item score point. An additional six to ten responses per item were chosen as part of the training pack. The anchor pack consisted of mid-range exemplars, while the training pack exemplars illustrated the range within each score point. The chief readers, who worked closely with QACs for each content area, facilitated the selection of response exemplars. One of the greatest difficulties in the selection of anchor and training exemplars was finding a sufficient number of papers representing the highest scores (4 or 8) as such scores are fairly rare.

Selecting and Training Scoring Staff

Selecting Quality Assurance Coordinators (QACs) and Verifiers

Because the read-behinds performed by the QACs and verifiers moderated the scoring process and thus maintained the integrity of the scores, individuals to fill those positions were selected for their accuracy. In addition, QACs, who train readers to score each item in their content areas, were selected for their ability to instruct and for their level of expertise in their content areas. For this reason, QACs typically are retired teachers who have demonstrated a high level of expertise in their respective disciplines. The ratio of QACs and verifiers to readers was approximately 1:11.

Training Quality Assurance Coordinators and Verifiers

To ensure that all QACs provided consistent training and feedback, the chief readers spent two days training and qualifying the QACs, and the QACs reviewed all items with the verifiers before scoring. In addition, QACs rotated among tables, supervising readers and reading behind verifiers, who in turn read behind a different table of readers each day.

Selecting Readers

Applicants were required to demonstrate their ability by participating in a preliminary scoring evaluation. The I-Score system enables Advanced Systems to efficiently measure a prospective reader's ability to score student responses accurately. After having participated in a training session, applicants were required to achieve at least 80% exact scoring agreement for a qualifying pack consisting of 20 responses to a predetermined item in their content area. Those 20 responses were randomly selected from a bank of approximately 150, all of which had been selected by QACs and approved by the chief readers and developers.

Training Readers

The QACs first applied the language of the scoring guide for an item to its anchor pack exemplars. Once discussion of the anchor pack had concluded, readers attempted to score the training pack exemplars correctly. The QACs then reviewed the training pack and answered any questions readers had before actual scoring began. With this system, two aspects of scoring efficiency are in conflict. First, in order to minimize training expense, it is desirable to train each reader on as few items as possible. Second, to prevent reader drift and to minimize retraining requirements, it is desirable to score a given item in a brief period of time. But the lower the number of unique items each reader scores, the greater the number of readers required to score that item quickly. To minimize that conflict, we divided each subject area's readers into two or more groups. On the first day of scoring, each group was trained to score a different item. When a group had completed all of an item's responses, those readers were trained on another item (or set).

SCORING ACTIVITIES

Student test booklets at grade level 4 and student response booklets at grade levels 8 and 11 were digitally scanned and scored on a file server for a dedicated, secure LAN. I-Score then distributed digital images of student responses to readers. Training and scoring took place over a period of approximately two weeks.

Items were randomly assigned to readers; thus, each item in a student's response booklet was more than likely scored by a different reader. By using the maximum possible number of readers for each student, the procedure effectively minimized error variance due to reader sampling. Matrix writing prompts, as well as all common and matrix-constructed and extended-response items were scored once with a 2% read-behind to ensure consistency among readers and accuracy of individual readers.

MONITORING READERS

After a reader scored a student response, I-Score determined whether that response should also be scored by another reader, scored by a QAC or verifier, or routed for special attention. QACs and verifiers used I-Score to produce daily reader accuracy and speed reports. QACs and verifiers were able to obtain current reader accuracy reports and speed reports on-line at any time.

SCORING THE WRITING

Maine teachers and administrators were recruited to score the common writing prompt at in-state scoring sessions that were held in Bangor and Gorham, Maine. Teachers who participated in the scoring process developed skills in holistic evaluation of writing using a rubric aligned with the standards outlined in the Maine *Learning Results*. Those skills could then be applied to writing instruction in the classrooms, and the scoring of writing also gave participants an opportunity to read the range of student writing produced at each grade and to connect their current teaching practices with the recommendations in the Maine *Learning Results*. Administrators who participated gained skills helpful in improving the teaching and evaluation of writing in their schools. Maine teachers' involvement in scoring also created a network of teachers who served as a resource to their local and state schools.

GENERAL SCORING GUIDES

SHORT-ANSWER ITEMS

Score Point	Description
2	▪ The student's response provides a complete and correct answer.
1	▪ The student's response is partially correct. ▪ The student's response may be incomplete or contain errors.
0	▪ The student's response is totally incorrect or too minimal to evaluate.
B	▪ Blank/no response.

OPEN RESPONSE ITEMS

Score Point	Description
4	▪ The student completes all important components of the task and communicates ideas clearly. ▪ The student demonstrates in-depth understanding of the relevant concepts and/or processes. ▪ When instructed to do so, the student chooses more efficient and/or sophisticated processes. ▪ When instructed to do so, the student offers insightful interpretations or extensions (e.g., generalizations, applications, and analogies).
3	▪ The student completes the most important components of the task and communicates clearly. ▪ The student demonstrates understanding of major concepts even though he/she overlooks or misunderstands some less important ideas or details.
2	▪ The student completes most important components of the task and communicates those clearly. ▪ The student demonstrates that there are gaps in his/her conceptual understanding.
1	▪ The student shows minimal understanding. ▪ The student addresses only a small portion of the required task(s).
0	▪ The student's response is totally incorrect or irrelevant.
B	▪ Blank/no response.

Stylistic & Rhetorical Aspects of Writing
Topic Idea Development

1	2	3	4	5	6
▪ Little topic development and/or organization, few details ▪ Possible evidence of voice ▪ Simplistic language (wording and sentence structures)	▪ Limited topic development, focus, and/or details ▪ Evidence of voice ▪ Limited variety in language used (wording and sentence structures)	▪ Moderate topic development, focus, and details ▪ Some voice ▪ Some variety in language used (wording and sentence structures)	▪ Well developed with control and relevant details ▪ Consistent voice ▪ Variety in language used (wording and sentence structures)	▪ Fully developed with strong details ▪ Sustained voice and/or tone with emerging style ▪ Effective use of language	▪ Topic and details richly developed ▪ Distinctive voice, tone and style ▪ Rich use of language

Analytic Annotations

		Commendations	Needs
Topic Development	The overall effect of the paper	TX sustained development throughout TY creative, insightful, and/or shows voice	TJ less repetition of ideas TK more development of ideas/topic
Organization	The degree to which the response is: ▪ Focused ▪ Clearly and logically ordered ▪ Clarified by paragraphs	OX clearly focused OY logical organization of ideas	OJ clearer focus OK more effective use of paragraphing
Details	The degree to which the response includes examples that develop the main points	DX details support focus DY uses interesting details	DJ to avoid simply listing details DK more/relevant details
Language/Style	The degree to which manipulation of language, including vocabulary, word choice, word combination, and sentence variety is effective	LX word choice enhances meaning LY sentence variety is used effectively	LJ more attention to word choice LK more variety in language

Standard English Conventions

1	2	3	4
▪ Errors seriously interfere with communication and/or ▪ Little control of sentence structure, grammar and usage, and mechanics in first writing draft	▪ Errors interfere somewhat with communication and/or ▪ Few or no errors in simplistic or limited text in first writing draft	▪ Errors do not interfere with communication and/or ▪ Few errors relative to length of essay or complexity of sentence structure, grammar and usage, and mechanics in first draft writing	▪ Control of a variety of sentence structures, grammar and usage, and mechanics ▪ Length and complexity of essay provide opportunity for student to show control of standard English conventions in first draft writing

Analytic Annotations

		Commendations	Needs
Sentences	The degree to which the response includes sentences that are correct in structure	SP correct sentence structure	SR correct sentence structure
Grammar and Usage	The degree to which the response demonstrates correct ▪ Use of standard grammatical rules of English ▪ Word usage and vocabulary	GUP correct application of grammatical rules GUQ control of vocabulary and word usage	GUR correct application of grammatical rules GUS greater attention to correct word usage
Mechanics	The degree to which the response demonstrates correct ▪ Punctuation ▪ Capitalization ▪ Spelling	MP control of mechanics aids clarity MQ correct mechanics in sophisticated construction	MR greater control of mechanics MS more careful editing

CHAPTER 12—EQUATING AND SCALING

Scaled scores for the 1999–00 MEA reading, writing, mathematics, science, and social studies tests were developed by equating the 1999–00 raw scores to the 1998–99 raw scores. Equating scores from alternate forms of a test adjusts for any difference in difficulty and allows for scores from the different forms to be comparable. Because the 1998–99 and 1999–00 versions of each test are developed from the same framework, they may be considered alternate forms. Equating test scores from the 1998–99 and 1999–00 administration of each test makes it possible to report the results of the 1999–00 administration on the same scale on which the MEA results were reported the previous year. Equating simply converts raw points from MEA 1999–00 to the MEA 1998–99 raw score scale. The equated scores then get translated to scaled scores. The process of equating and scaling does not change the rank ordering of students, give more weight to particular questions, or change students' performance-level classifications.

Equating for MEA used the **anchor-test-nonequivalent-groups** design with external anchor described by Petersen, Kolen, and Hoover (1993). The anchor test for reading, mathematics, science, and social studies was a set of matrix items included in both test administrations. These items were external to the test in that they did not contribute to the students' raw scores in either administration of the test. The groups of students who took each test in 1998–99 and 1999–00 were naturally occurring groups and no assumption was made regarding their equivalence. Item response theory (IRT) is particularly useful in this type of equating (Allen & Yen, 1979). All IRT calibrations performed on MEA were for equating.

Equating for MEA writing used the reading scaled scores as the anchor test. The Tucker Method described in Kolen and Brennan (1995, pp. 105–111) was implemented.

Developing equated scores for the 1999–00 MEA involved several steps. The first step was to construct the anchor test; that is, to determine the set of equating items. The second step was to calibrate the items in an IRT model. The IRT model used was a combination of the three-parameter logistic (3PL) model for multiple-choice items, the two-parameter logistic (2PL) model for short-answer items, and the graded response model (GRM) for the open-response items. The calibration was first performed on the 1998–99 data. The item parameters of the equating items resulting

from this calibration were fixed for the calibration of the 1999–00 data. Fixing the parameters of the equating items ensured that the two "forms" of the test (1998–99 and 1999–00) were calibrated to the same scale of the trait being measured. Using test characteristic curves (TCC), raw scores from the 1999–00 MEA were mapped or equated to raw scores on the 1998–99 MEA. The equated scores were then translated to the 500 to 580 scale. The following sections detail this equating process.

DETERMINING THE SETS OF EQUATING ITEMS

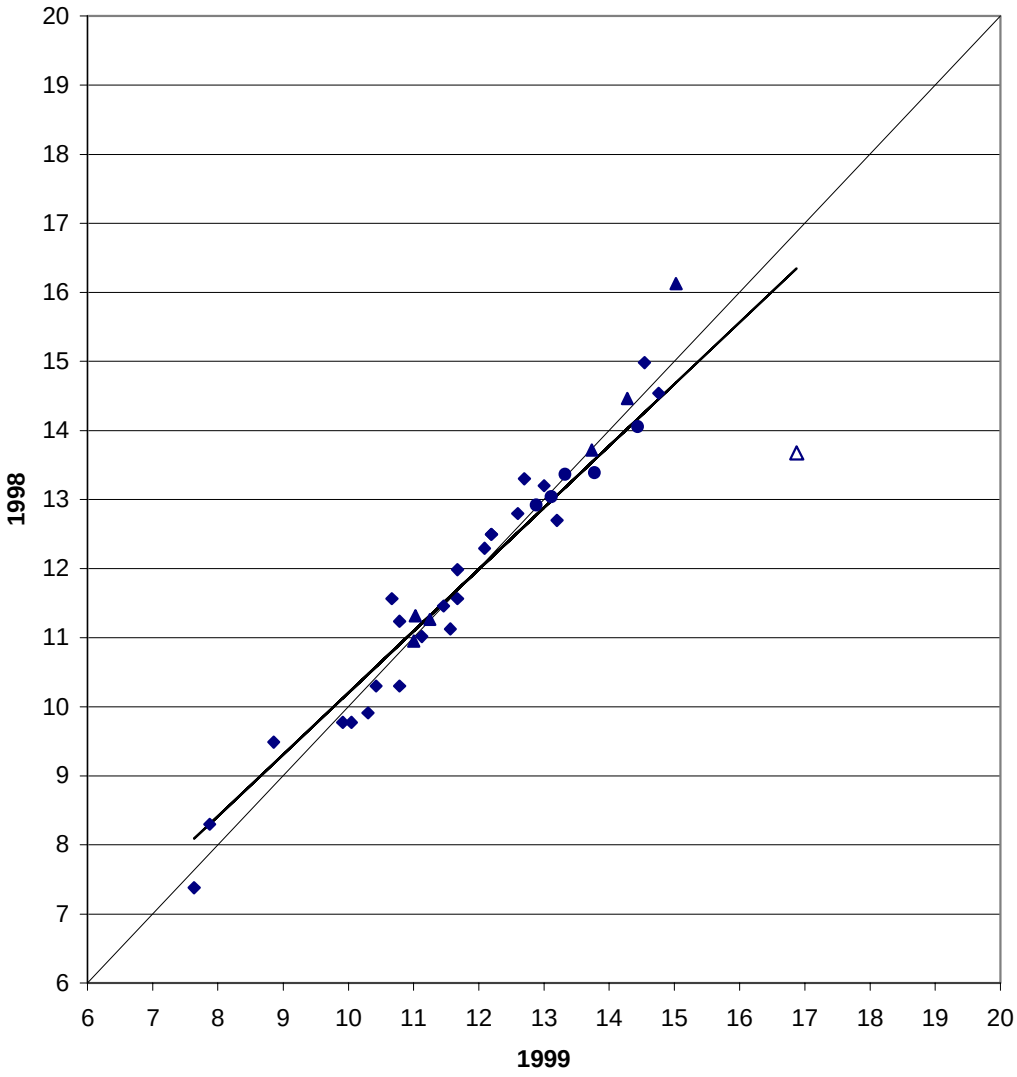
During the development stage of the 1999–00 MEA, matrix items that were also administered in 1998–99 were identified as potential equating items. These items were designated based on the following criteria:

1. The average difficulty of the equating items is about the same as the average difficulty of the 1998–99 test.
2. The total points from the equating items are about equivalent to 40 percent of the total points on the test.
3. The position of each item in the 1999–00 form is about the same as its position in the 1998–99 form.
4. The distribution of the items across different relevant categories (i.e., items types and content areas) is similar to that of the whole test.
5. There should not be any significant change in the item from one administration to the other.

To determine the final set of equating items for each grade level and subject combination a differential item functioning (DIF) approach using the delta plot method was applied. The p-values of each multiple-choice and short-answer item were transformed to the delta metric. Each item had two p-values, one for each test administration. The delta scale is an inverse normal transformation of percentage correct to a linear scale with a mean of thirteen and standard deviation of four (Holland & Wainer, 1993). A high delta value indicates a difficult item. For open-response items, the average score divided by the maximum possible score or adjusted p-value were transformed to the delta metric. The delta values computed for the potential equating items was plotted for each subject (reading, mathematics, science, and social studies) in each grade level (4, 8, 11).

Figure 12–1 is an example of a delta plot for equating items. The dark diagonal line is the trend line and the light diagonal line is the identity line. Different shapes were used to identify different item types: for multiple-choice items; for short-answer items; and for open-response items. The perpendicular distance of each item to the regression line was computed. The unshaded shape indicates the item with the greatest perpendicular distance from the regression line. Items that were not more than three standard deviations away from the regression line were used as equating items.

Figure 12–1
Sample Delta Plot



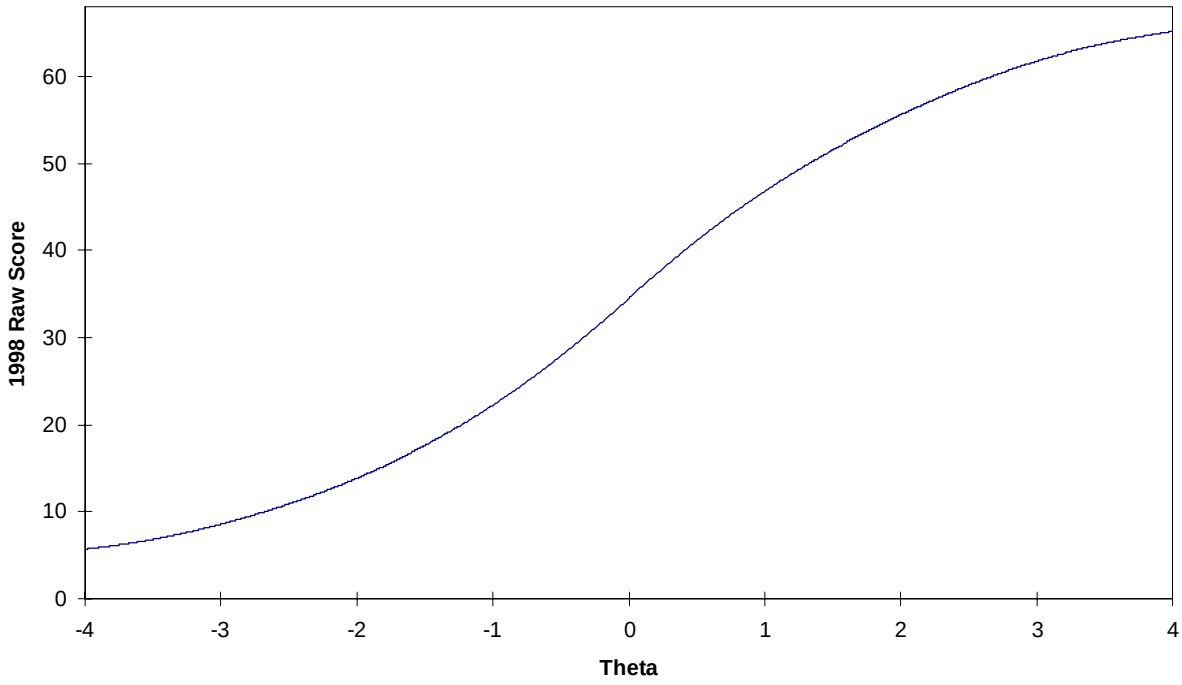
(MC SA OR)

ITEM CALIBRATIONS

IRT calibration was performed on the common and matrix items from the 1998–99 MEA using a combination of IRT models specific to item types (i.e., 3PL for multiple-choice, 2PL for short-answer, and GRM for open-response). Each of these models expresses examinees, tendencies to achieve certain scores on the items contributing to a scale as a function of a parameter that is not directly observed and commonly referred to a θ. Using the current version of PARSCALE, item parameters were estimated based on those models. From the parameter estimates, a test characteristic curve (TCC) was obtained using common items only—the same set of items on which individual student scores for the 1998–99 MEA were based. Through this TCC, each raw score on the test can be mapped to a unique value of θ. An example of a TCC is shown in Figure 12–2.

Figure 12–2
Sample Test Characteristic Curve

Grade 4 English Language Arts



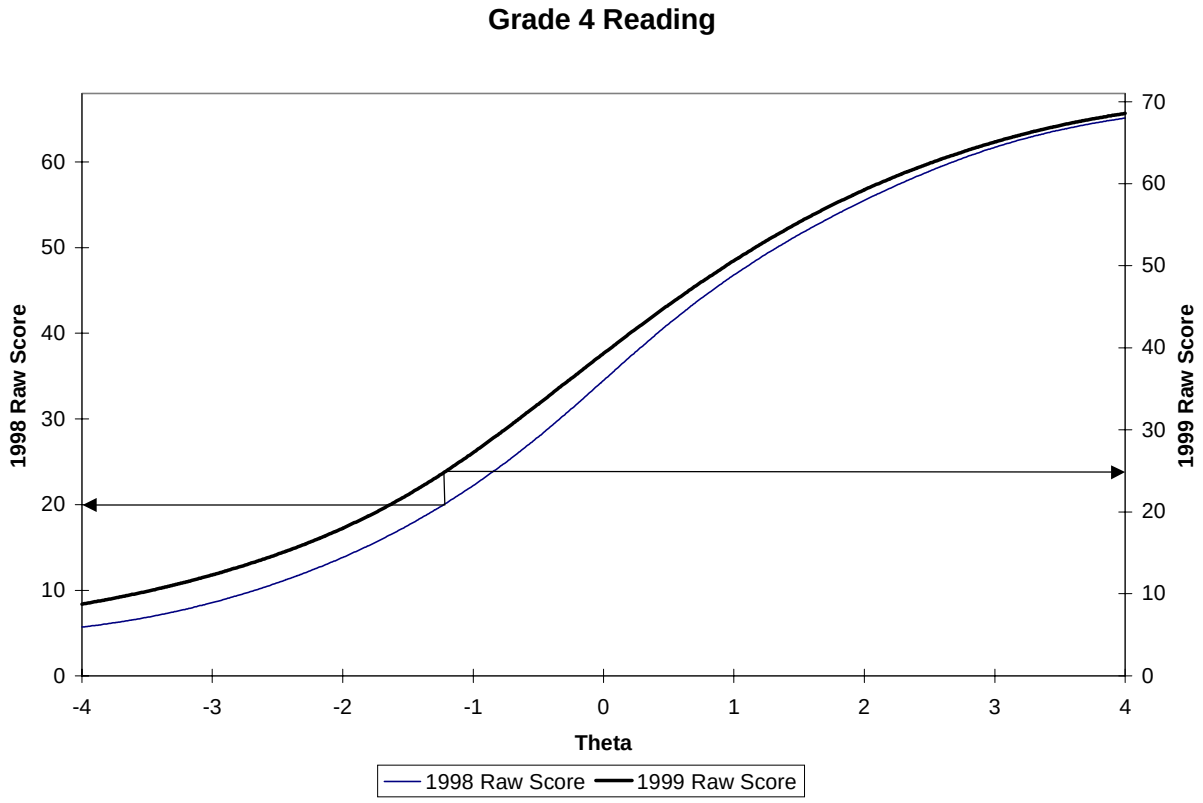
An IRT calibration was also performed on the 1999–00 MEA student response data. This data set included responses to 1999–00 MEA common and matrix items. So that 1999–00 MEA would be calibrated to the same θ scale as 1998–99 MEA, IRT parameters for the equating items were not estimated for this calibration. Instead, they were fixed to the estimated values resulting from the calibration of the 1998–99 data.

EQUATED SCORES FOR READING AND WRITING

The item parameter estimates for the common items were used to obtain the TCC for 1999–00 reading test. Using this TCC, each raw score can be mapped to a θ value. Because the TCCs for the 1998–99 and 1999–00 MEAs were on the same θ metric, for each value of θ there is a corresponding raw score for each of the 1998–99 and 1999–00 common item sets. Thus, for each grade in reading, each 1999–00 raw score can be mapped to a 1998–99 raw score. For example, using the TCCs in Figure 12-3, a raw score of 25 in 1999–00 maps to a raw score of 20 in 1998–99. This mapping is essentially IRT true-score equating (Lord, 1980) using the fixed-b method to maintain a consistent θ metric.

For writing, the Tucker Method was used so that each 1999–00 raw score was mapped to a 1998–99 raw score.

Figure 12-3
Finding Equated Scores



SCALED SCORES FOR READING AND WRITING

After raw scores from the 1999–00 reading and writing tests were mapped to 1998–99 raw scores (i.e., equated scores), these scores were translated to scaled scores. The functions that translate raw scores to scaled scores are

$$S = m_1 r + b_1 \quad \text{if } r < P, \text{ and}$$
$$S = m_2 r + b_2 \quad \text{if } r > P$$

where S is the scaled score, r is the raw score, and P is the threshold for "meets the standards." The scaling parameters m_1, m_2, b_1, and b_2 are based on the results of standard-setting processes implemented for reading and writing in 1998–99.

Linear scaling parameters were determined so the minimum scaled score for "partially meets the standards" was 521, the minimum scaled score for "meets the standards" was 541, and the minimum scaled score for "exceeds the standards" was 561. This was done by solving two linear equations relating the raw threshold scores to these predetermined scaled-score values. The values for *m*s and *b*s for reading and writing are in Table 12-1.

After the transformation constants were applied, scores were rounded to the nearest even integer. Transformed scores below 502 were reported as 502; transformed scores above 580 were reported as 580.

SCALED SCORES FOR HEALTH AND VISUAL AND PERFORMING ARTS

The equating procedure for health and visual and performing arts was the same as that for reading, mathematics, science, and social studies. However, the scaled scores for health and visual and performing arts are linear transformations of θ scores, not raw scores like in reading, mathematics, science, and social studies. The functions that translate θs to scaled scores are

$$S = m_1\theta + b_1 \quad \text{if } \theta < P, \text{ and}$$
$$S = m_2\theta + b_2 \quad \text{if } \theta > P$$

where S is the scaled score, θ is the ability estimate, and P is the threshold for "meets the standards." These scaling parameters m_1, m_2, b_1, and b_2 are based on the results of standard-setting processes implemented for health and visual and performing arts in 1998–99. These constants are also presented in Table 12-1.

Table 12-1
Transformation Constants Used to Compute Scaled Scores for Reading, Writing, Health, and Visual and Performing Arts

Grade	Subject Area	Transformation Constants			
		m_1	b_1	m_2	b_2
4	Reading	1.55	488.66	1.61	486.70
	Writing	2.47	495.08	2.31	498.11
	Health	19.68	533.95	10.13	537.37
	Visual and Performing Arts	8.21	534.14	11.40	531.48
8	Reading	1.69	484.95	1.67	485.63
	Writing	2.19	501.32	2.79	490.60
	Health	12.29	537.45	10.74	537.89
	Visual and Performing Arts	9.39	534.99	14.29	531.86
11	Reading	1.84	472.59	1.45	486.90
	Writing	2.92	482.21	2.49	490.85
	Health	13.89	536.26	10.78	537.32
	Visual and Performing Arts	5.12	536.29	14.81	527.37

SCORES FOR MATHEMATICS, SCIENCE, AND SOCIAL STUDIES

For mathematics, science, and social studies, IRT parameters resulting from the calibrations were used to estimate student abilities. The estimated student abilities were based only on common items. The cumulative distributions of raw scores and scaled scores for each subject and grade combination for 1999–00 and 1998–99 were used to find the equated cutpoints. Thus, for 1999–00, a new set of cutpoints was obtained. This process is described in Figure 12–4. Suppose $c_{1998–99}$ is a cutpoint resulting from the standard setting in 1998–99. This cutpoint is in the raw score metric. Using the frequency distribution of the raw scores for 1998–99 the cumulative percentage associated with this cutpoint was estimated through linear interpolation. Using the frequency distribution of ability estimates, the θ value associated with this cumulative percentage was determined. Because ability for 1998–99 and 1999–00 are on the same θ scale, the obtained θ value corresponds to the same ability for both years. The 1999–00 cumulative percentage associated with this θ was then mapped to a 1999–00 raw score through linear interpolation, resulting in $c_{1999–00}$.

The above process was used for each cutpoint set in 1998–99 for each grade for mathematics, science, and social studies. The resulting curpoints are presented in Table 12-2. These cutpoints were used to obtain new scaling

parameters m_1, m_2, b_1, and b_2 which were then used to compute the scaled scores for 1999–00. The new scaling parameters are presented in Table 12-3. The functions that translate raw scores to scaled scores are

$$S = m_1 r + b_1 \quad \text{if } r < P, \text{ and}$$
$$S = m_2 r + b_2 \quad \text{if } r > P$$

where S is the scaled score, r is the raw score, and P is the threshold for "meets the standards."

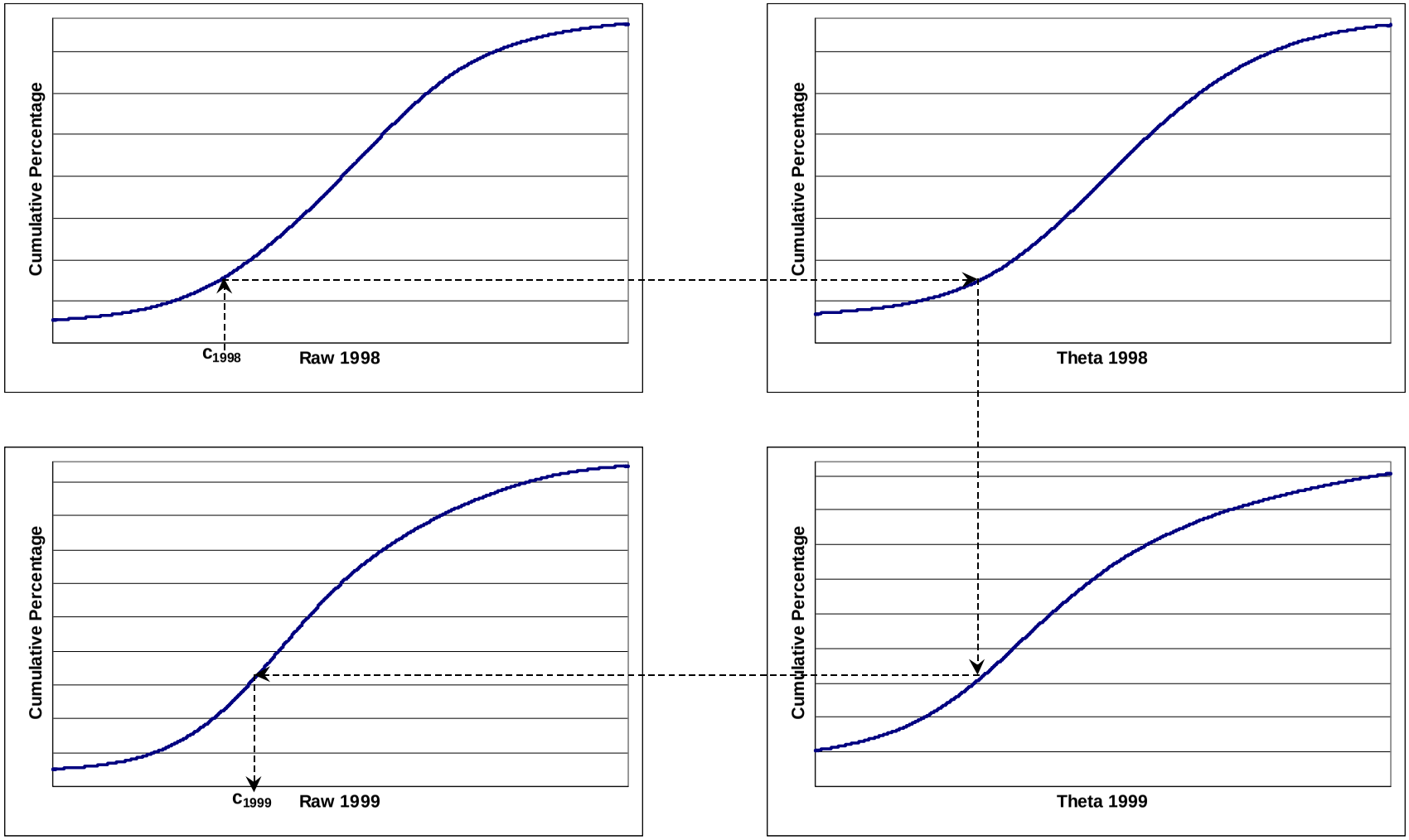
Table 12-2
Threshold (Minimum) Total Test Score for Each Performance Category for Mathematics, Science, and Social Studies

Grade	Subject Area	Maximum Score on Test	Threshold Score		
			Exceeds the Standards	Meets the Standards	Partially Meets the Standards
4	Mathematics	50	43.89	32.58	19.87
	Science	50	43.25	37.32	20.14
	Social Studies	50	39.32	28.08	16.41
8	Mathematics	50	44.76	31.98	18.44
	Science	50	41.12	31.76	18.69
	Social Studies	50	39.36	29.83	17.50
11	Mathematics	50	42.89	28.96	14.36
	Science	50	41.75	31.55	16.08
	Social Studies	50	39.36	27.08	17.23

Table 12-3
Transformation Constants Used to Compute Scaled Scores for mathematics, science, and social studies

Grade	Subject Area	Transformation Constants			
		m_1	b_1	m_2	b_2
4	Mathematics	1.77	483.40	1.57	489.74
	Science	3.37	415.13	1.16	497.57
	Social Studies	1.78	491.05	1.71	492.86
8	Mathematics	1.57	490.93	1.48	493.78
	Science	2.13	473.23	1.53	492.40
	Social Studies	2.09	478.41	1.62	492.61
11	Mathematics	1.43	499.40	1.37	501.34
	Science	1.96	479.13	1.29	500.21
	Social Studies	1.63	496.88	2.03	486.05

Figure 12–4
Finding Equated Cutpoints



CHAPTER 13—ITEM ANALYSES

As noted in Brown (1983), "a test is only as good as the items it contains." A complete evaluation of a test's quality must include an evaluation of each question. Both the *Standards for Educational and Psychological Testing* and the *Code of Fair Testing Practices in Education* include standards for identifying quality questions. Questions should assess only knowledge or skills that are under assessment and should avoid assessing irrelevant factors. They should also be unambiguous and free of grammatical errors, potentially insensitive content or language, and other confounding characteristics. Further, questions must not unfairly disadvantage test takers from particular racial, ethnic, or gender groups.

Both qualitative and quantitative analyses were conducted to ensure that MEA questions met these standards. Previous sections in this report have delineated the qualitative checks on question quality. This chapter focuses on more quantitative evaluations. The statistical evaluations are presented in three sections: (1) difficulty indices, (2) item-test correlations, and (3) subgroup differences in question performance. The results presented in this chapter are based on the statewide administration of the MEA in March of 1998–99. About 17,000 grade 4 students, 18,000 grade 8 students, and 15,000 grade 11 students participated in the assessment.

DIFFICULTY INDICES

All multiple-choice, short-answer, and open-response questions were evaluated in terms of difficulty and relationship to overall score according to standard classical test theory practice. Difficulty was measured by averaging the proportion of points received across all students who received the question. Multiple-choice and short-answer questions were scored dichotomously (correct versus incorrect), so for these questions, the difficulty index was simply the proportion of students who correctly answered the question. Open-response questions allowed for scores between 0 and 4. By computing the difficulty index as the average proportion of points received, the indices for multiple-choice, short-answer, and open-response questions were placed on a similar scale; the index ranged from 0 to 1 regardless of the question type. Although this index is traditionally described as a measure of difficulty (as it is described here), it is properly interpreted as an easiness index because larger values indicate easier questions.

An index of 0 indicates that no student received credit for the question, and an index of 1 indicates that every student received full credit for the question.

Questions that are answered correctly by almost all students provide little information about differences in student ability, but they do indicate knowledge or skills that have been mastered by most students. Similarly, questions that are correctly answered by very few students may indicate knowledge or skills that have not yet been mastered by most students, but such questions provide little information about differences in student ability. In general, to provide best measurement, difficulty indices should range from near-chance performance (.25 for four-option, multiple-choice questions or essentially zero for short-answer and open-response questions) to .90. Indices outside this range indicate questions that were either too difficult or too easy for the target population.

Although difficulty is an important question characteristic, the relationship between performance on a question and performance on the whole test or a relevant test section may be more critical. A question that assesses relevant knowledge or skills should relate to other questions that are purported to be measuring the same knowledge or skills.

ITEM-TEST CORRELATIONS

Within classical test theory, these relationships are assessed using correlation coefficients that are typically described as either item-test correlations or, more commonly, discrimination indices. The discrimination index used to analyze MEA multiple-choice items and zero- or one-scored short-answer items was the point-biserial correlation between item score and a criterion total score on the test. As such, the index ranged from –1 to 1, with the magnitude and sign of the index indicating the relationship's strength and direction, respectively. For open-response items, item discrimination indices were based on the Pearson product-moment correlation. The theoretical range of these statistics was also from –1 to 1, with a typical range from .3 to .6.

In general, discrimination indices are interpreted as indicating the degree to which high- and low-ability students perform differently on a question or, equivalently, the degree to which performance on a question helps to differentiate between high- and low-ability students. From this perspective, indices near 1 indicate that high-ability students are more likely to answer the question correctly, indices near –1 indicate that low-ability students are more

likely to answer the question correctly, and indices near 0 indicate that performance on the question is equally likely to be answered correctly by high- and low-ability students.

Discrimination indices can be thought of as measures of how closely a question assesses the same knowledge and skills assessed by other questions contributing to the criterion total score; that is, the discrimination index can be interpreted as a measure of construct consistency. In light of this interpretation, the selection of an appropriate criterion total score is crucial to the interpretation of the discrimination index. For the 1999–2000 MEA, the criterion score for each common item was the total score for all common items. For each matrix item the criterion score was the total score for the form that item belonged to.

SUMMARY OF ITEM ANALYSIS RESULTS

Frequency distributions and summary statistics of the difficulty and discrimination indices for each question are provided in Tables 13-1 through 13-2. In general, the question difficulty and discrimination indices are in acceptable and expected ranges. Very few questions were answered correctly at near-chance or near-perfect rates. Similarly, the positive discrimination indices indicate that most questions were assessing consistent constructs, and students who performed well on individual questions tended to perform well overall. There were a small number of questions with near-zero discrimination indices, but none were reliably negative. Occasionally, questions with less desirable statistical characteristics need to be included in assessments to ensure that content is appropriately covered, but there were very few such cases on the MEA.

A comparison of indices across grade levels is complicated because these indices are population dependent. Direct comparisons would require that either the questions or the students were common across groups. However, one can say that with respect to multiple-choice items, the fourth- and eighth-grade students tended to have more difficulty answering the mathematics questions on the fourth- and eighth-grade tests as compared to the eleventh-grade students answering the math questions on the eleventh-grade tests. In science, the opposite of this statement is true in that eleventh-grade students had more difficulty answering multiple-choice science questions on the eleventh-grade tests than did the other two grades in their respective science tests. The fourth-grade students may have had a slightly easier time with the reading questions on the fourth-grade tests as compared to the eighth-grade students taking the

eighth-grade reading questions. Similarly, eighth-grade students may have had a slightly easier time with the reading questions on the eighth-grade tests as compared to the eleventh-grade students taking the eleventh-grade reading questions.

Comparisons within grade levels are reasonable with one caveat: in comparing common and matrix questions, one assumes that the sampling scheme for matrix questions ensures that the students receiving a particular matrix question are representative of the entire population that received the common questions. With that caveat in mind, there appeared to be immaterial differences in the mean difficulty of common and matrix multiple-choice questions regardless of grade level and subject. The exceptions to this observation were the social studies multiple-choice questions for grade 11 (with a mean difference of 0.16).

Comparing the difficulty indices of multiple-choice and short-answer or open-response questions is inappropriate because multiple-choice questions can be answered correctly by guessing. Thus, it is not surprising that the difficulty indices for multiple-choice questions tend to be higher (indicating easier questions) than the difficulty indices for other question types. Similarly, the partial credit allowed by open-response questions is advantageous in the computation of question-test correlations, so the discrimination indices for these questions tend to be larger than the discrimination indices of other question types.

Table 13-1
Average Difficulty and Discrimination of Different Question Types for Each Subject: Grade 4

Subject	Statistics	Multiple-Choice			Short-Answer	Constructed-Response
		Common	Matrix	All		
Reading	*n*	18	72	90	29	15
	Average Difficulty	.65	.62	.63	.55	.39
	Average Discrimination	.44	.46	.45	.52	.61
Mathematics	*n*	21	59	80	29	17
	Average Difficulty	.54	.56	.55	.43	.38
	Average Discrimination	.41	.38	.39	.49	.57
Science	*n*	20	72	92	17	17
	Average Difficulty	.59	.63	.62	.52	.31
	Average Discrimination	.34	.36	.36	.37	.46
Social Studies	*n*	20	72	92	17	17
	Average Difficulty	.67	.55	.58	.54	.30
	Average Discrimination	.38	.33	.34	.42	.49
Health	*n*		72		12	36
	Average Difficulty		.64		.65	.45
	Average Discrimination		.32		.35	.46
Visual and Performing Arts	*n*		72			12
	Average Difficulty		.56			.44
	Average Discrimination		.30			.38

Table 13-2
Average Difficulty and Discrimination of Different Question Types for Each Subject: Grade 8

Subject	Statistics	Multiple-Choice			Short-Answer	Constructed-Response	Extended-Response
		Common	Matrix	All			
Reading	*n*	18	72	90	29	15	1
	Average Difficulty	.61	.65	.64	.59	.46	.41
	Average Discrimination	.41	.43	.42	.53	.65	.68
Mathematics	*n*	20	60	80	29	15	1
	Average Difficulty	.55	.51	.52	.34	.30	.27
	Average Discrimination	.43	.40	.41	.53	.63	.68
Science	*n*	20	72	92	17	17	
	Average Difficulty	.57	.53	.54	.45	.28	
	Average Discrimination	.40	.32	.34	.40	.52	
Social Studies	*n*	20	72	92	17	17	
	Average Difficulty	.61	.55	.56	.36	.33	
	Average Discrimination	.41	.36	.37	.40	.59	
Health	*n*		72		12	12	12
	Average Difficulty		.65		.69	.46	.42
	Average Discrimination		.34		.43	.50	.53
Visual and Performing Arts	*n*		72			12	
	Average Difficulty		.57			.39	
	Average Discrimination		.33			.44	

Table 13-3
Average Difficulty and Discrimination of Different Question Types for Each Subject: Grade 11

Subject	Statistics	Multiple-Choice			Short-Answer	Constructed-Response	Extended-Response
		Common	Matrix	All			
Reading	*n*	18	72	90	29	15	1
	Average Difficulty	.72	.67	.68	.64	.51	.48
	Average Discrimination	.49	.48	.48	.58	.69	.73
Mathematics	*n*	20	60	80	29	15	1
	Average Difficulty	.44	.41	.42	.26	.28	.31
	Average Discrimination	.40	.39	.39	.52	.68	.73
Science	*n*	20	72	92	17	17	
	Average Difficulty	.51	.48	.49	.39	.31	
	Average Discrimination	.36	.34	.35	.41	.58	
Social Studies	*n*	20	72	92	17	15	1
	Average Difficulty	.51	.35	.39	.27	.35	.41
	Average Discrimination	.38	.24	.27	.39	.61	.66
Health	*n*		72		12	17	9
	Average Difficulty		.71		.71	.50	.47
	Average Discrimination		.43		.46	.59	.57
Visual and Performing Arts	*n*		72			12	
	Average Difficulty		.55			.42	
	Average Discrimination		.36			.50	

SUBGROUP DIFFERENCES IN QUESTION PERFORMANCE

The *Code of Fair Testing Practices in Education* explicitly states that subgroup differences in performance should be examined when sample sizes permit, and actions should be taken to make certain that differences in performance are due to construct-relevant, rather than irrelevant, factors. The *Standards for Educational and Psychological Testing* includes similar guidelines. As part of the effort to identify such problems, MEA questions were evaluated in terms of differential item functioning (DIF) statistics.

DIF procedures are designed to identify questions for which subgroups of interest perform differently beyond the impact of differences in overall achievement. For the MEA, the standardization DIF procedure (Dorans & Kulick, 1986) was employed to evaluate subgroup differences between male and female students. This procedure calculates the difference in item performance for groups of students matched for achievement on the total test. That is, the average item performance is calculated for students at every total score, then an overall average is calculated weighting the total score distribution so it is the same for the two groups.

The index ranges from –1 to 1 for multiple-choice and short-answer questions and is adjusted to the same scale for open-response questions. Negative numbers indicate that the question was more difficult for female students. Positive numbers indicate that the question was easier for female students.

Dorans and Holland (1993) suggested that index values between –0.05 and 0.05 should be considered negligible for dichotomously scored questions (such as MEA multiple-choice questions). Most MEA questions fall within this range. Dorans and Holland further stated that dichotomously scored questions with values between –0.10 and –0.05 and between 0.05 and 0.10 (i.e., low DIF) should be inspected to ensure that no possible effect is overlooked, and that questions with values outside the [–0.10, 0.10] range (i.e., high DIF) are more unusual and should be examined very carefully. These standards can be applied to open-response questions by accounting for the larger range of possible index values and scaling appropriately. That is, values of the DIF index can range from –4.0 to 4.0, so the

corresponding ranges are between –0.2 and 0.2 for negligible difference, between –0.4 and –0.2 and between 0.2 and 0.4 for low DIF and outside [-0.4, 0.4] for high DIF.

DIF indices indicate differential performance between two groups. That differential performance may or may not be indicative of bias in the test. Course-taking patterns, group differences in interests, or differences in school curricula can lead to DIF. If subgroup differences in performance are related to construct-relevant factors, the questions should be considered for inclusion on a test.

Each question was categorized according to the guidelines adapted from Dorans and Holland (1993). Tables 13-5 to 13-7 provide the number of questions in each of the three DIF categories for male versus female for each grade level tested. There are some MEA questions categorized as low or high DIF. These indices must not be interpreted as indisputable evidence of bias. Both the *Code of Fair Testing Practices in Education* and the *Standards for Educational and Psychological Testing* assert that test questions must be free from construct-irrelevant sources of differential difficulty. If subgroup differences in performance can be plausibly attributed to construct-relevant factors, the questions may be included on a test. What is important is to determine if the cause of this differential performance is construct relevant.

Table 13-4

Differential Item Functioning (DIF) Categorization Item Type: Grade 4

SUBJECT	Item Type	High DIF				Low DIF				Negligible DIF			
		Favor Female	Favor Male	N	%	Favor Female	Favor Male	N	%	Favor Female	Favor Male	N	%
READING	Multiple-Choice	0	0	0	0.00	3	26	29	34.52	12	31	43	51.2
	Open-Ended	0	0	0	0.00	3	0	3	3.57	9	0	9	10.7
MATHEMATICS	Multiple-Choice	0	1	1	0.79	5	12	17	13.49	26	36	62	49.2
	Open-Ended	0	1	1	0.79	3	1	4	3.17	19	22	41	32.5
SCIENCE	Multiple-Choice	0	4	4	3.17	12	15	27	21.43	26	35	61	48.4
	Open-Ended	0	1	1	0.79	2	1	3	2.38	23	7	30	23.8
SOCIAL STUDIES	Multiple-Choice	0	1	1	0.79	3	14	17	13.49	42	32	74	58.7
	Open-Ended	0	0	0	0.00	2	0	2	1.59	16	16	32	25.4

Table 13-5

Differential Item Functioning (DIF) Categorization Item Type: Grade 8

SUBJECT	Item Type	High DIF				Low DIF				Negligible DIF			
		Favor Female	Favor Male	N	%	Favor Female	Favor Male	N	%	Favor Female	Favor Male	N	%
READING	Multiple-Choice	1	5	6	7.14	3	18	21	25.00	16	29	45	53.6
	Open-Ended	0	0	0	0.00	5	0	5	5.95	7	0	7	8.3
MATHEMATICS	Multiple-Choice	0	2	2	1.60	4	12	16	12.80	27	35	62	49.6
	Open-Ended	4	0	4	3.20	4	1	5	4.00	26	10	36	28.8
SCIENCE	Multiple-Choice	0	5	5	3.97	6	19	25	19.84	31	31	62	49.2
	Open-Ended	1	0	1	0.79	4	2	6	4.76	18	9	27	21.4
SOCIAL STUDIES	Multiple-Choice	0	0	0	0.00	4	18	22	17.46	27	43	70	55.6
	Open-Ended	0	2	2	1.59	9	2	11	8.73	8	13	21	16.7

Table 13-6

Differential Item Functioning (DIF) Categorization Item Type: Grade 11

SUBJECT	Item Type	High DIF				Low DIF				Negligible DIF			
		Favor Female	Favor Male	N	%	Favor Female	Favor Male	N	%	Favor Female	Favor Male	N	%
READING	Multiple-Choice	2	10	12	14.29	5	15	20	23.81	18	22	40	47.6
	Open-Ended	0	0	0	0.00	7	0	7	8.33	4	1	5	6.0
MATHEMATICS	Multiple-Choice	1	1	2	1.60	4	18	22	17.60	23	33	56	44.8
	Open-Ended	1	0	1	0.80	9	2	11	8.80	26	7	33	26.4
SCIENCE	Multiple-Choice	0	6	6	4.76	6	26	32	25.40	16	38	54	42.9
	Open-Ended	0	0	0	0.00	6	2	8	6.35	14	12	26	20.6
SOCIAL STUDIES	Multiple-Choice	0	5	5	4.00	7	17	24	19.20	22	41	63	50.4
	Open-Ended	2	1	3	2.40	10	6	16	12.80	6	8	14	11.2

CHAPTER 14—RELIABILITY

Although an individual question's performance is an important focus for evaluation, a complete evaluation of an assessment must also address the way that questions function together and complement one another. Any measurement includes some amount of measurement error; that is, no measurement can be perfectly accurate. This is true of academic assessments—no assessment can measure students perfectly accurately; some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. Questions that function well together produce assessments that have less measurement error; that is, the errors made should be small on average. Such assessments are described as reliable.

There are a number of ways to estimate an assessment's reliability. One approach is to split all test questions into two groups and then correlate students' scores on the two half tests. This is known as a split-half estimate of reliability. If the two half-test scores correlate highly, questions on the two half tests must be measuring very similar knowledge or skills. This is evidence that the questions complement one another and function well as a group. This also suggests that measurement error will be minimal.

The split-half method requires the psychometrician to select which questions contribute to each half-test score. This decision may have an impact on the resulting correlation. Cronbach (1951) provided a statistic that avoids this concern about the split-half method. Cronbach's coefficient is an estimate of the average of all possible split-half reliability coefficients.

RELIABILITY AND STANDARD ERRORS OF MEASUREMENT

Table 14-1 presents descriptive statistics, Cronbach's coefficient, and raw- and scaled-score standard errors of measurement for each subject separately for each grade level. The reported reliability for writing, health, and visual and performing arts are the averages of the computed Cronbach's across forms. The low reliability values can be attributed to the lower number of items in each form in those tests.

Note: two scaled-score standard errors of measurement are presented: one for scaled scores below 542 and one for scaled scores of 542 and above. This is because different slopes are used in the linear transformation to scaled scores at these two parts of the scaled-score range.

Table 14-1
Reliabilities, Standard Errors of Measurement, and Descriptive Statistics
MEA 1999–2000

Grade	Subject	n	Raw Score						Scaled Score	
									<542	>=542
			Min.	Max.	Mean	S.D.	Rel.	S.E.M.	S.E.M.	S.E.M.
4	Reading	15,517	1	46	25.52	7.03	.81	3.06	1.96	1.59
	Writing	15,402	4	30	15.32	4.21	.53	2.89	4.92	2.70
	Mathematics	15,945	0	50	25.21	8.99	.85	3.48	2.53	1.39
	Science	16,062	2	47	24.19	7.08	.77	3.40	3.20	0.88
	Social Studies	16,044	2	48	24.30	7.31	.80	3.27	2.39	1.48
	Health*	16,112	1	19	10.95	2.73	.51	1.91	3.93	5.04
	Visual and Performing Arts*	16,747	0	10	5.13	2.16	.56	1.43	7.18	5.95
8	Reading	16,641	0	47	27.25	7.16	.83	2.95	2.03	1.41
	Writing	16,557	4	30	16.92	4.43	.62	2.72	4.79	2.65
	Mathematics	16,871	0	50	22.13	10.37	.88	3.59	2.56	1.38
	Science	16,969	0	48	22.95	8.41	.84	3.36	2.68	1.14
	Social Studies	16,945	0	47	24.29	8.17	.85	3.16	2.36	1.24
	Health*	17,024	1	20	10.94	3.10	.48	2.24	4.57	3.40
	Visual and Performing Arts*	17,757	0	10	5.03	2.20	.61	1.37	7.79	5.90
11	Reading	13,982	2	48	32.39	5.99	.79	2.74	2.21	1.18
	Writing	13,845	4	30	17.66	4.62	.63	2.82	4.75	4.08
	Mathematics	13,720	0	50	19.10	10.08	.84	4.03	2.71	1.69
	Science	13,842	0	47	21.21	7.48	.81	3.26	2.71	1.17
	Social Studies	13,849	0	49	21.91	8.17	.82	3.47	2.38	1.64
	Health*	14,140	1	20	12.13	3.09	.45	2.29	5.43	3.90
	Visual and Performing Arts*	14,753	0	10	5.02	2.37	.63	1.44	7.69	3.46

*The reported reliability is the average reliability across forms.

STRATIFIED COEFFICIENT

According to Feldt and Brennan (1989), a prescribed distribution of items over categories (such as different item types) indicates the presumption that at least a small, but important, degree of unique variance is associated with the categories. In contrast, Cronbach's coefficient is built upon the assumption that there are no such local or clustered dependencies. A stratified version of coefficient corrects for this problem.

Stratified coefficient was calculated separately for each common item test and grade level. The stratification was based on item types (multiple-choice versus open-response). These results are provided in Table 14-2.

Table 14-2
Coefficients and Stratified
MEA 1999-2000

Grade	Content		$_{mc}$	N_{mc}	$_{or}$	N_{or} (Pts.)	Stratified
4	Reading	0.812	.682	18	.736	9 (30)	0.824
	Mathematics	0.852	.771	20	.754	10 (30)	0.860
	Social Studies	0.803	.681	20	.720	10 (30)	0.814
	Science	0.768	.637	20	.662	10 (30)	0.776
8	Reading	0.827	.708	18	.757	9 (30)	0.838
	Mathematics	0.880	.795	20	.832	9 (30)	0.895
	Social Studies	0.849	.736	20	.787	10 (30)	0.860
	Science	0.836	.757	20	.735	10 (30)	0.845
11	Reading	0.790	.631	18	.730	9 (30)	0.805
	Mathematics	0.838	.761	20	.749	9 (30)	0.852
	Social Studies	0.819	.704	20	.746	9 (30)	0.833
	Science	0.807	.643	20	.755	10 (30)	0.821

RELIABILITY OF PERFORMANCE-LEVEL CATEGORIZATION

All test scores contain measurement error; thus, classifications based on test scores are also subject to measurement error. After the performance levels were specified and students were classified into those levels, empirical analyses were conducted to determine the statistical accuracy and consistency of the classifications.

ACCURACY

Accuracy refers to the extent to which decisions based on test scores match decisions that would have been made if the scores did not contain any measurement error. Accuracy must be estimated because errorless test scores do not exist.

CONSISTENCY

Consistency measures the extent to which classification decisions based on test scores match the decisions based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test questions if two complete, parallel forms of the test are given to the same group of students. This is usually impractical, especially on lengthy tests such as the MEA. To overcome this issue, techniques have been developed to estimate both accuracy and consistency of classification decisions based on a single administration of a test. The technique developed by Livingston and Lewis (1995) was used for the MEA because it can be used with both constructed-response and multiple-choice questions.

CALCULATING ACCURACY

All of the accuracy and consistency estimation techniques described below make use of the concept of "true scores" in the sense of classical test theory. A true score is the score that would be obtained on a test that had no measurement error. It is a theoretical concept that cannot be observed, although it can be estimated. Following Livingston and Lewis (1995), the true-score distribution for the MEA was estimated using a four-parameter beta distribution, which is a flexible model that allows for extreme degrees of skew in test scores.

In the Livingston and Lewis method, the estimated true scores are used to classify students into their "true" performance category, which is labeled "true status." After various technical adjustments (which are described in Livingston & Lewis, 1995), a 4×4 contingency table is created for each test and grade level. The cells in the table are the proportion of students who were classified into each performance category by the actual (or observed) scores on the MEA (i.e., observed status) and by the true scores (i.e., true status). As an example, Table 14-3 shows the

accuracy contingency table for fourth-grade social studies. The accuracy contingency tables for all grades and subjects are provided in Appendix A (under step 5). Additional steps in the analysis are also shown in Appendix A.

Table 14-3
Accuracy Contingency Table for Grade 4 Social Studies

True Status	Observed Status			
	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards
Does Not Meet the Standards	**.10**	.03	.00	.00
Partially Meets the Standards	.05	**.46**	.07	.00
Meets the Standards	.00	.06	**.21**	.01
Exceeds the Standards	.00	.00	.00	**.01**

Proportions on the diagonal (in bold) indicate exact agreement between the observed status and true status. If the test were perfectly accurate, all of the off-diagonal cells would be zero. Accuracy is the sum of the diagonal (i.e., the proportion of exact agreement across the four performance levels). In Table 14-3, the diagonal sums to .78, indicating that 78 percent of the students were classified into exactly the same performance categories by their observed scores and their true scores.

KAPPA

Another way to measure consistency is to use Cohen's (1960) coefficient κ (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. Cohen's κ can be used to estimate the classification consistency of a test from two parallel forms of the test. The second form in this case was the one estimated using the Livingston and Lewis (1995) method. Cohen's κ is shown in Table 14-4. Because κ is corrected for chance, the values of κ are lower than the other consistency estimates in Table 14-4.

CALCULATING CONSISTENCY

To estimate consistency, the true scores are used to estimate the distribution of classifications on an independent, parallel test form. After statistical adjustments (see Livingston and Lewis, 1995) are made, a new 4×4 contingency table is created for each test and grade level that shows the proportions of students who were classified into each

performance category by the actual test and by another (hypothetical) parallel test form. Consistency, which is the proportion of students classified into exactly the same categories by the two forms of the test, is the sum of the diagonal for the new contingency table. The consistency contingency tables are shown under step 7 in Appendix A.

RESULTS OF ACCURACY, CONSISTENCY, AND KAPPA ANALYSES

The accuracy, consistency, and kappa indices for all grades and subjects are summarized in Table 14-4.

Table 14-4
Estimates of Accuracy and Consistency of Performance-Level Classification

Grade	Subject	Accuracy	Consistency	Kappa (κ)
4	Reading	0.78	0.69	0.48
	Writing	0.86	0.81	0.57
	Mathematics	0.79	0.7	0.53
	Science	0.83	0.75	0.50
	Social Studies	0.78	0.69	0.47
	Health	0.77	0.66	0.27
	Visual and Performing Arts	0.55	0.44	0.19
8	Reading	0.79	0.71	0.50
	Writing	0.86	0.8	0.63
	Mathematics	0.81	0.74	0.59
	Science	0.79	0.71	0.53
	Social Studies	0.78	0.70	0.52
	Health	0.75	0.65	0.29
	Visual and Performing Arts	0.58	0.48	0.25
11	Reading	0.77	0.68	0.46
	Writing	0.84	0.77	0.61
	Mathematics	0.79	0.71	0.55
	Science	0.82	0.74	0.52
	Social Studies	0.75	0.66	0.48
	Health	0.72	0.61	0.23
	Visual and Performing Arts	0.65	0.53	0.30

For certain decisions, concern may be highest regarding decisions made about a particular threshold. For example, if a college gave credit to students who achieved an Advanced Placement test score of 4 or 5, but not 1, 2, or 3, one might be interested in the accuracy of the dichotomous decision below 4 versus 4 or above. Table 14-5 reports accuracy and consistency for various dichotomous categorizations on the MEA. MEA P/M cut accuracy ranges from

.77 to .97 and M/E accuracy ranges from .97 to .999. These are relatively high values compared to the 1999–00 Advanced Placement accuracy of decisions based on the 2–3 cut and 3–4 cut, which ranges from .84 to .95.

Table 14-5

Accuracy and Consistency of Dichotomous Categorizations

Grade	Subject	Accuracy			Consistency		
		D/P	P/M	M/E	D/P	P/M	M/E
4	Reading	.93	.86	.98	.91	.81	.97
	Writing	.94	.92	.99+	.91	.90	.99+
	Mathematics	.90	.90	.98	.85	.86	.98
	Science	.85	.97	.99+	.79	.96	.99+
	Social Studies	.92	.87	.99	.88	.82	.98
	Health	.98	.81	.98	.95	.72	.98
	Visual and Performing Arts	.79	.78	.96	.70	.70	.94
8	Reading	.94	.87	.98	.92	.81	.97
	Writing	.95	.91	.99+	.92	.88	.99+
	Mathematics	.91	.92	.99	.87	.88	.98
	Science	.89	.91	.99	.85	.87	.99
	Social Studies	.91	.89	.99	.87	.84	.98
	Health	.97	.79	.99+	.95	.71	.99
	Visual and Performing Arts	.82	.80	.95	.73	.73	.91
11	Reading	.95	.85	.98	.93	.79	.96
	Writing	.95	.90	.98	.93	.86	.98
	Mathematics	.88	.92	.99	.84	.88	.99
	Science	.87	.95	.99+	.82	.93	.99+
	Social Studies	.89	.88	.98	.84	.83	.97
	Health	.95	.78	.99	.92	.69	.98
	Visual and Performing Arts	.79	.86	.98	.71	.81	.95

CHAPTER 15—VALIDITY

As noted in the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1985, p. 9), "validity is the most important consideration in test evaluation." Validity refers to whether specific inferences made from test scores are appropriate, meaningful, and useful. There are several types of validity-related evidence that can be used to support appropriate, meaningful, and useful inferences based on test scores.

CONTENT-RELATED EVIDENCE

As noted in the *Standards* (p. 10), evidence of test validity begins with test development and continues throughout the entire testing process. Chapters 2 through 9 provide evidence regarding the alignment between the content of the MEA and Maine's *Learning Results*.

EXTERNAL EVIDENCE

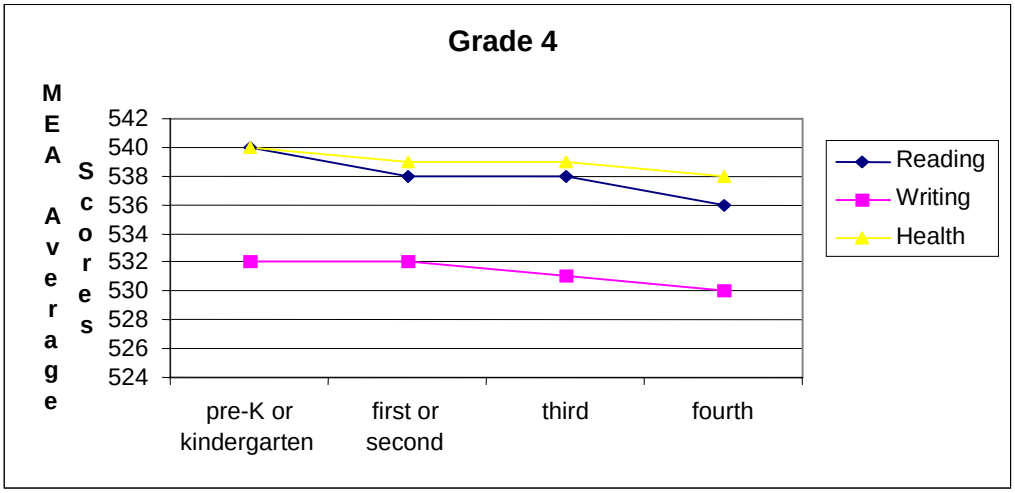
External validity of the MEA is conveyed by the relationship between test scores and situational variables such as school transience, course-taking pattern, attitude toward subject matter, and self-image. These situational variables were all based on student questionnaire data collected during the administration of the MEA. Note that not all the questionnaire items referred to in the following subsections were asked regarding all of the subjects assessed by the MEA. Note also that no inferential statistics are included. However, because the numbers of students were large enough, differences in average scores could be shown to be statistically significant.

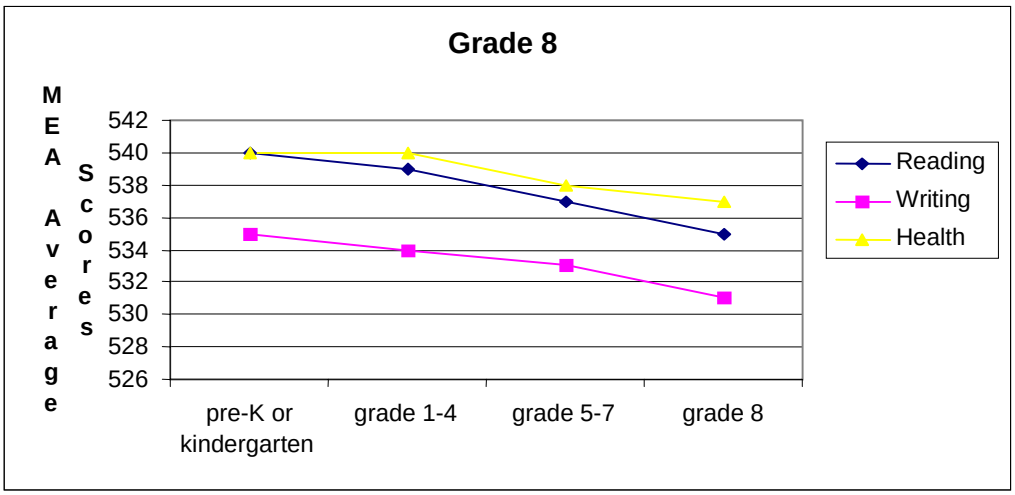
SCHOOL TRANSIENCE

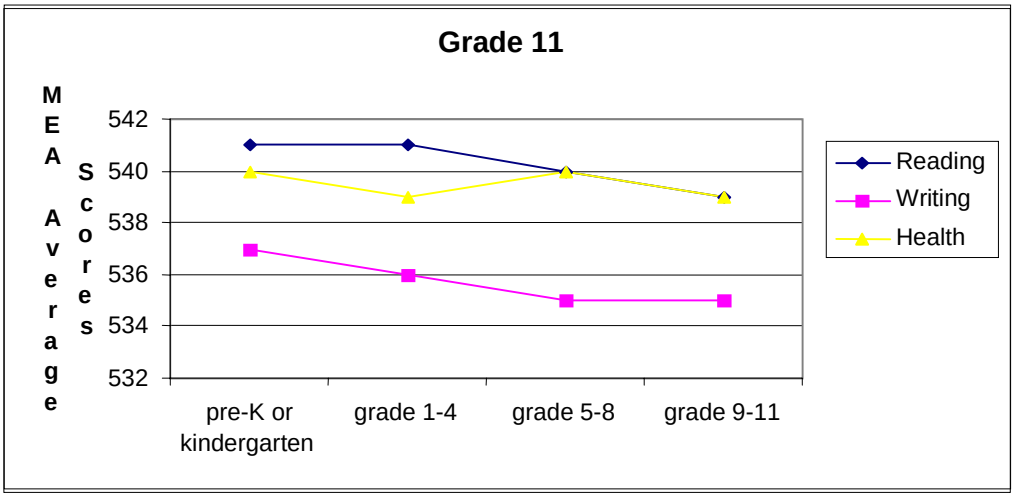
This is an evaluation of how time in a single school is related to test scores. Students were asked, "In what grade did you start coming to school in this school district?" Medsker (1998) found that typically, students who change schools often do not perform as well as students who regularly attend a single school or school system. Charts in Figure 15-1 clearly indicate that students who spent more time in a single school tended to have higher test scores in reading, science, and visual and performing arts.

Figure 15-1
School Transience and MEA Scores

Question: In what grade did you start coming to school in this school district?





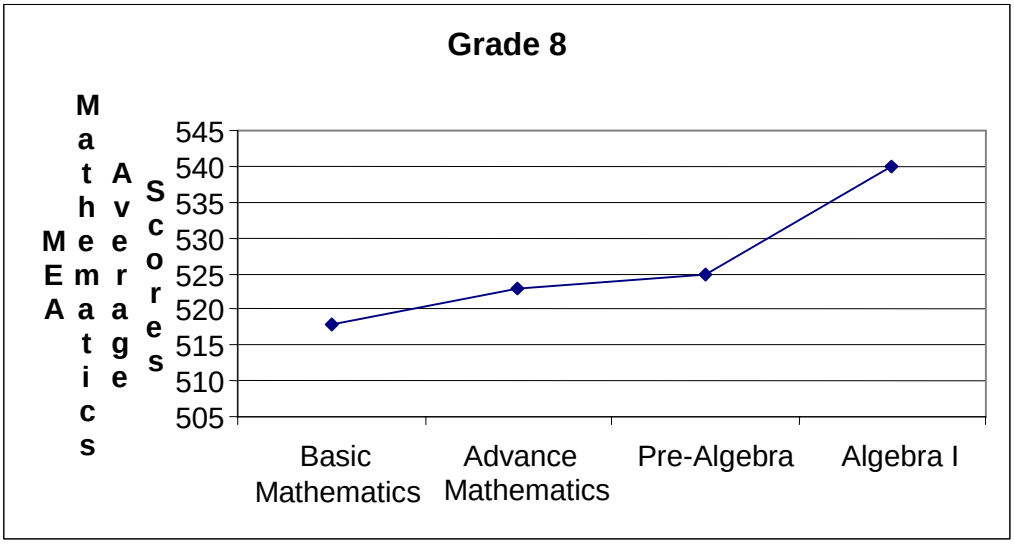


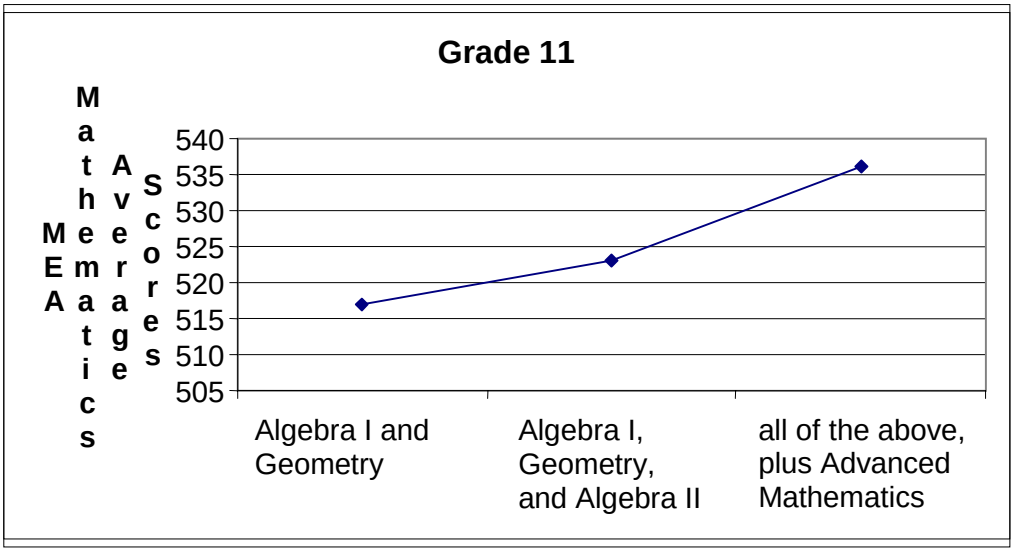
COURSE-TAKING PATTERN

Grades 8 and 11 examinees were asked questions related to their course-taking patterns in mathematics. Eighth graders were asked, "What best describes the mathematics class you are taking in the eighth grade?" and eleventh graders were asked, "What best describes the mathematics courses will you complete before you graduate?" The charts in Figure 15-2 both show that the higher-level mathematics courses are associated with higher MEA mathematics scores.

Figure 15-2
MEA Mathematics Scores and Course-Taking Patterns

Grade 8 Question: What best describes the mathematics class you are taking in the eighth-grade?
Grade 11 Question: What best describes the mathematics courses will you complete before you graduate?



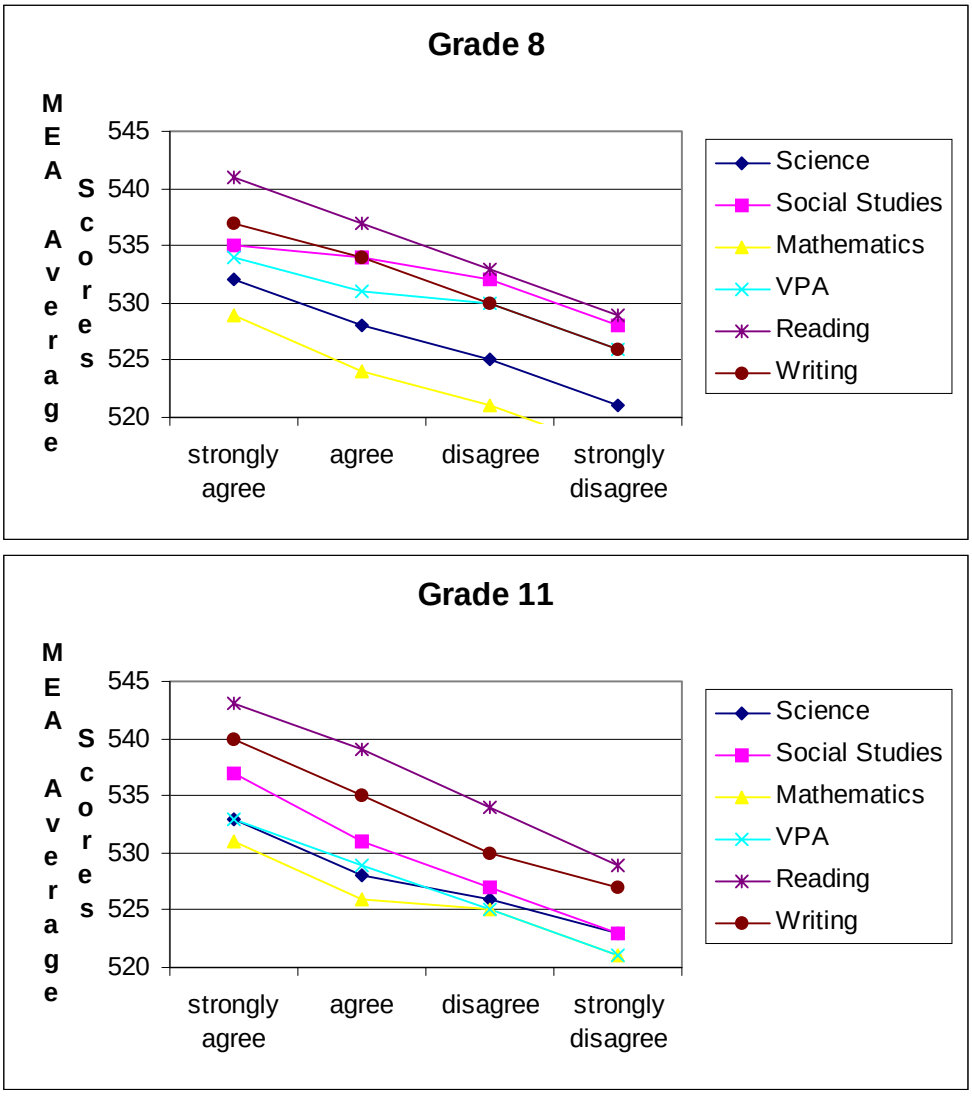


Questionnaire items related to examinees' attitudes toward different subjects tested in the MEA were administered to eighth and eleventh graders. For reading, mathematics, science, social studies, and visual and performing arts, students were asked how they felt about the statement, "My knowledge of [subject] will be useful to me in my future work." For health, students were asked how they feel about the statement, "My knowledge about health education will be helpful to me as an adult." Charts in Figure 15-3 show that students' degree of agreement with statements that indicate their attitudes toward the subjects tested in the MEA are related positively to MEA scores.

ATTITUDE TOWARD SUBJECT MATTERS

Figure 15-3
Attitude Toward Subject Matters and MEA Scores

Question: My knowledge of [subject] will be useful to me [in my future work/as an adult].

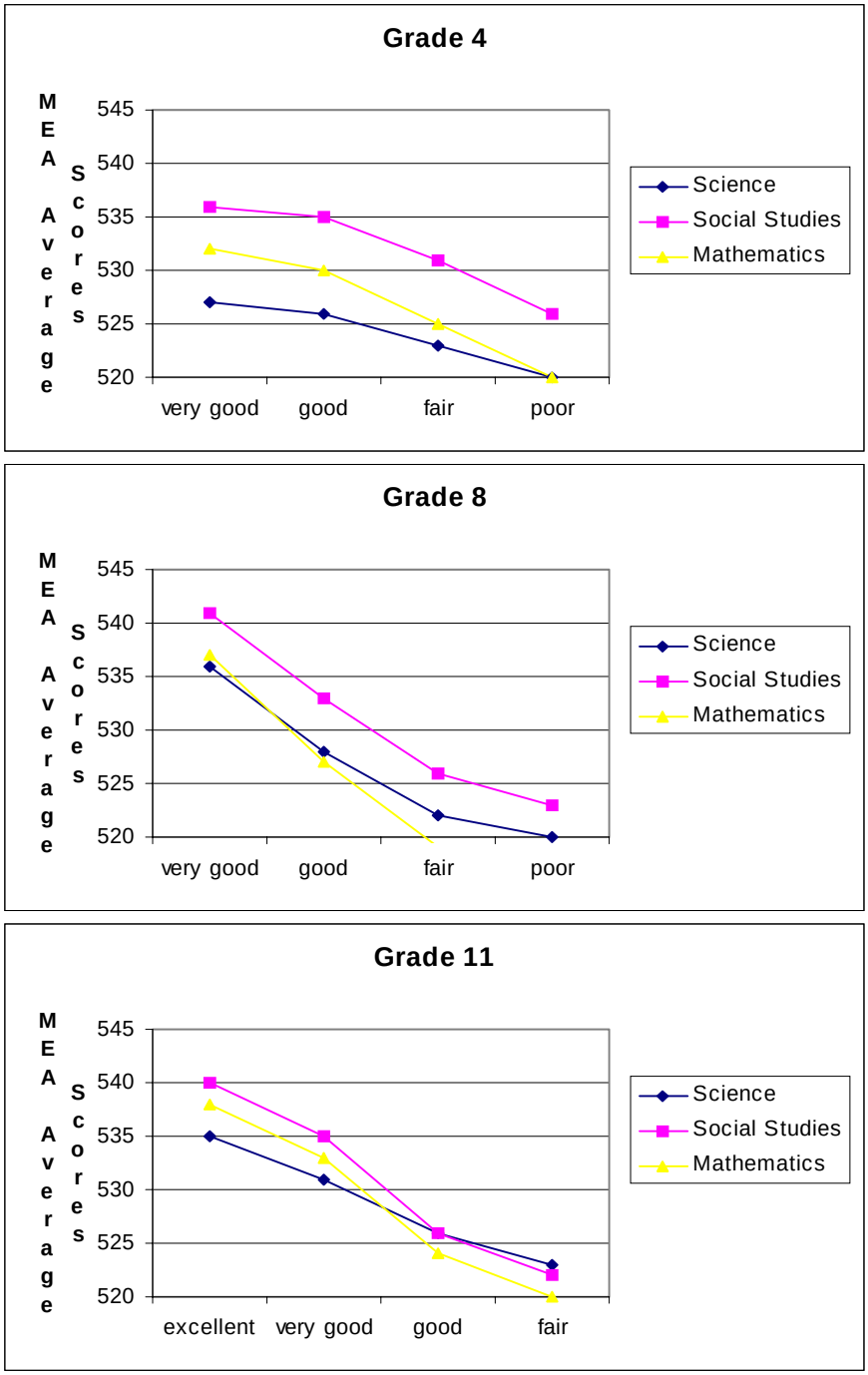


SELF-IMAGE

All students who participated in the MEA were asked, "What best describes how you are as a student?" Figure 15-4 indicates that there is a positive relationship between students' self-image and their MEA scores in mathematics, social studies, and health.

Figure 15-4
Self-Image and MEA Scores

Question: What best describes how you are as a student?



CHAPTER 16—SCORE REPORTING

Table 16-1 lists the primary MEA reports.

Table 16-1 Primary MEA Reports	
1.	*Student Report for Parents/Guardians*
2.	*Student Labels*
3.	*School Common Item Level Class Report*
4.	*School Report*
5.	*District Report*
6.	*Student Writing CD*

STUDENT REPORT FOR PARENTS/GUARDIANS

Student reports show the scaled score for each subject area, as well as a score band that indicates the standard error of measurement surrounding each score. Performance level definitions are provided so that parents/guardians will understand how to interpret the scaled scores. Specific comments are provided about the student's writing performance. Information is also provided to show how the student's performance compared to the average scores from the student's school, district, and state. An overview of test content is provided, along with a cautionary statement about interpreting scores and guidelines for parents/guardians for helping their children improve.

STUDENT LABELS

To aid schools in keeping track of student scores, schools were supplied with student score information on individual labels that they could affix to files, if desired.

SCHOOL COMMON ITEM LEVEL CLASS REPORT

The *Common Item Level Class Report* shows the answers that each student gave on the multiple-choice questions, as well as his/her score on each open-response question. The report also summarizes overall performance at the school, district, and state levels for each of the question types.

SCHOOL AND DISTRICT REPORTS

The school and district reports are intended for administrators and other interested parties. The school report includes performance level definitions, scaled score intervals, and information about how summary statistics are affected by students not tested; all of which are intended to help the reader interpret the report. The school report provides all results for the school, the district, and the entire state. The results provided are

- the number of students tested by student status (regular, students with disabilities, and limited English proficient students) for all subject areas combined and separately for each subject area,
- the percentage of students in each performance level by subject area,
- the distribution of scaled scores by subject area,
- the number of students in each performance level by subject area and student status,
- subject area subscores outlining the number of possible points by learning results standards,
- three-year comparisons of school results, and
- average subject score by number of years in the school or district.

The district report is the same as the school report, except that it does not include the school-level data and the three-year comparisons are by district rather than by school.

STUDENT WRITING CD

The student writing CD contains all of the student's writing for each school. The schools are then able to print out and/or review the actual student's work.

Sample reports can be found in Appendix B.

QUALITY CONTROL (QC) PROCESS FOR ENSURING ACCURACY OF PRINTED REPORTS

GENERAL

1. Whenever new reports are received from Measurement, Design, and Analysis (MDA), the *date and time* they were received is written at the top of each report so that it will be easy to identify the most recent version of each report.
2. For each of the items that follows, a checkmark was put in a logical position on each report to indicate that each check was done. For instance, after verifying that a name is correct, a checkmark is placed next to the name; after verifying that a score is correct or that a bar is the correct length, a checkmark is placed next to it; and so on. This lets other QC staff verify which checks have been done and which have not.
3. When all checks are completed on a given report, the QC staff's initials and that day's date are written at the top of the page so that everyone knows who checked them.

PARENT REPORTS:

Letter Side:

1. Proofing text and formatting of entire side is done once thoroughly, and then spot-checked in additional QC runs.
2. The State MEA Summary Results (bottom right box): the percentages are verified that they match those on the school and district reports for the state (page 2 bar graph, page 4, and page 6). The bars are then checked to make sure that they accurately represent the percentages reported.

Performance Assessment Side:

1. Proofing text and formatting of entire side is done once thoroughly, and then spot-checked in additional QC runs.
2. It is verified that the student name and grade are the same as those printed on the letter side.
3. QC staff also checks to make sure that the performance level corresponds to the scaled score.
4. They also make sure that the diamond placement in the top box corresponds to score and performance levels and that the range bar does not fall outside of the scale area.

5. If the student was excluded or testing was incomplete, it was verified that no scaled score or performance level appeared, nor were there diamonds or range bars. Instead, it was verified that the words "*excluded*" or "*testing incomplete*" appeared in the performance level box.
6. The performance level and scaled score was compared to the common item report to ensure that they matched. They were also compared to the labels to ensure that they matched.
7. It was verified that the school, district, and state averages matched those in the school and district reports (page 2), and they also verified the accuracy of the height of the bars. To make sure that the height of the bar reflected the number on top of the bar, QC staff looked to the left of the bars at the scale. (The bar height should match the performance level.)
8. It was verified that there were no student bars if a student was excluded or testing was incomplete. (Instead, he/she would get the school, district, and state bars only.)
9. Writing comments were checked that the commendations/needs correspond to the comment codes on the Common Item Class Report (for individual students). It was also checked that the comments were properly categorized (e.g., needs statements into Needs box and commendations statements into Commendation box).
10. Students marked as **NT** (not tested) or **TI** (tested incomplete) *may* still have comments. It was verified that any comments matched what was reported on the Common Item Level Report.
11. Student's Performance in Content Area subcategories: Diamond placement was verified. It was checked that the diamonds did *not* overlap borders, nor did their corners get cut off. It was also checked that there was a diamond for each of the three categories. If a student was *excluded* or the testing was *incomplete*, then there were *no* diamonds. (If a student had scores for Writing but was incomplete or excluded for Reading, then the diagram would show two diamonds in the Writing category, but no diamond for Reading.)

LABELS

1. Spelling, punctuation, and formatting (for margins, fit of text on the label, and so on) were checked.
2. It was verified that the school and district information is correct.
3. The names, proficiency levels, and scaled scores were checked to make sure they matched what was reported on the common item and parent reports.
4. It was also verified that the students listed as belonging to a given school were the same on all reports for that school.
5. The grade was verified and also that each page of labels includes information for only *one* school.

COMMON ITEM CLASS REPORTS (READING AND WRITING)

1. The QC staff was directed to proof the text and formatting of the report, including the legend (on reverse side), if provided.
2. They also compared the heading information to the shells and verified that the data in the heading matched the data in the shells.
3. They then verified that the names appear in alphabetical order, and in groups of five.
4. The staff was then told to **highlight** the information for any student who was *excluded or incomplete* (marked with **asterisks**). Subtract these students from the total and indicate the new total next to the original "group size" indicated in the box at the top of the page (this is the number used when calculating averages).
5. It was verified that the number of points per score did not exceed the maximum value indicated in the heading. (If the number 4 is written in the total possible points box then no one should have an 8 for a score.)
6. The keys were then verified by comparing each correct answer to the incorrect answers listed underneath for each question. (For example, if A is the correct MC, there should be no A's for incorrect answers.)
7. Next, the number of students receiving each *type* of annotation was counted. A need or commendation with the same 1st letter should only be counted once per student. (For example, a student who received two *needs* that began with a T [for "Topic Development"] and one *commendation* that began with a T, would only be counted *once* for the *needs* and *once* for the *commendation*.) These numbers should match those reported on page 8 of the school and district reports.
8. Then the QC staff calculates the average scaled score and the average points earned for the school.

9. Finally, the match to school and district reports takes place by adding across classes to get school scores, and across schools to get district scores (remember to skip the highlighted students and divide by the adjusted group size).

$$\frac{\text{Total of all scaled scores}}{\text{Total number of students}} = \text{average scaled score for the class}$$

$$\frac{\text{Total of all points earned}}{\text{Total number of students}} = \text{average points earned for the class}$$

SCHOOL AND DISTRICT REPORTS

Page 1:

1. The entire page was proofed for both text and formatting errors, including verifying the page references in the table of contents.

Page 2:

1. The entire page was proofed for both text and formatting errors once thoroughly.
2. It was verified that the scaled scores matched the ones on the parent report and the state-score handout (provided by MDA).
3. The percentage tables were then checked to make sure that the state percentages matched those on the parent reports and handout. The school and district should match the percentages on page 4, 6, or 9.
4. The scores reported for the school and district under Average Performance Score to the averages calculated from the common item reports were compared next.
5. Then the staff calculated and verified the accuracy of the Cum. Avg. under Average Performance Score. Total both averages for last year and this year and divide by two.
6. Finally, they compared this last year's reports to verify historic data.

Page 3:

1. The entire page was proofed for both text and formatting errors. The informational paragraph at the top of the page was checked so that it refers to school or district as appropriate.
2. It was verified that the students enrolled on (school report and district report) equaled the number(s) listed as group size on the common item report.

3. All percentages (except the last two rows) were computed by taking the number in each row and dividing it by the number enrolled.

Pages 4, 6, and 9 (Reading and Writing) and pages 4, 6, 8, and 11 (Mathematics, Science & Tech., and Social Studies, and VPA):

1. All pages were proofread for both text and formatting errors once thoroughly.
2. QC staff added up the number of students at each performance level (school "N" and district "N") to get the total included for that content area. And it was verified that it matched the number of students on the common item report (the modified total, minus excluded and incomplete students).
3. Then the percent of students at each level was verified by dividing the number at that level by the total number of students included for that content area. Add the percents down the levels to make sure they equal 99-101.
4. Under "Average Points Attained," the percentage for school, district, and state was verified by dividing the number ("N") by the number of points possible. (Note: If the school or district is small, *some of these cells may be blank*. Each Learning Results Content Standard must have *at least 5 students* to be reported in this table.)

Pages 5, 7, and 10 (Reading and Writing) and pages 5, 7, 9, and 11 (Mathematics, Science & Tech., Social Studies, and VPA):

1. All pages were proofread for both text and formatting errors once thoroughly.
2. It was verified that the percentages for each option equaled 99-101 per question.
3. QC staff then checked percentages for reasonableness. (If the total *number of students* in a category is *less than 5*, no percentage will be reported. Percentages for "special" categories, such as "Migrant," might total less than 100.)

Page 8: Summary of Annotations Table (Writing Only):

1. The entire page was proofread for both text and formatting errors once thoroughly.
2. It was verified that the number of students receiving a commendation or need matched the number counted on the common item report. A need or commendation with the same 1st letter was only counted once per student.

[For example, a student who received two *needs* that began with a T (for "Topic Development") and one *commendation* that began with a T, would only be counted *once* for the *needs* and *once* for the *commendation*.]

3. Staff then recalculated the percentages by dividing the number of students reported in this table by the total number of students tested in writing for the school and/or district. To get the number tested in writing, add up the number of students at each performance level on page six.

SECTION IV: REFERENCES

American Educational Research Association, American Psychological Association & National Council on Measurement in Education (1985). *Standards for educational and psychological testing.* Washington, DC: APA

Brown, F.G. (1983). *Principles of educational and psychological testing* (3rd Edition). Fort Worth: Holt, Rinehart and Winston

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334

Dorans, N.J., & Holland, P.W. (1993). DIF detection and description. In P.W. Holland & H. Wainer (Eds.) *Differential item functioning* (pp. 35–66). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Feldt, L.S. & Brennan, R.L. (1989). Reliability. In R.L. Linn (Ed.) *Educational Measurement* (3rd Ed., pp. 105–146).

Kingston, N.M., Kahl, S.R., Sweeney, K.P., & Bay, L. (2001). Setting performance standards using the Body of Work (BoW) method. In G.J. Cizek (Ed.) *Setting Performance Standards: Concepts, Methods, and Perspectives* (pp. 219–248). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kolen, M.J., & Brennan, R.L. (1995). *Test equating methods and practices.* New York: Springer-Verlag.

Livingston, S.A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement*, 32, 179–197.

Livingston, S.A., & Zieky, M.J. (1982). *Passing scores.* Princeton, NJ: Educational Testing Service.

Lord, F.M. (1980). A*pplications of item response theory to practical testing problems.* Hillsdale, NJ: Erlbaum.

Medsker, G.J. (1998). *Determining the relationship between student transience and KIRIS school results: Are schools with transient students unfairly impacted?* (HumRRO Report, FR-WATSD-98-12). Alexandria, VA: Human Resources Research Organization.

Petersen, N.S., Kolen, M.J., & Hoover, H.D. (1989) Scaling, norming, and equating. In R.L. Linn (Ed.), *Educational measurement* (3rd Ed., pp. 221-262).

SECTION V: APPENDICES

APPENDIX A

1999-2000 ACCURACY AND CONSISTENCY OF CLASSIFICATION TABLES

Accuracy and Consistency of Classifications

Grade 4 Reading

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.019558	0.00921	0.00000	0.000000	‖	0.02877
Partially Meets the Standards	0.060806	0.49774	0.05605	0.000002	‖	0.61450
Meets the Standards	0.000011	0.06908	0.24207	0.010725	‖	0.32184
Exceeds the Standards	0.000000	0.00000	0.00942	0.025284	‖	0.03471
	0.080376	0.57603	0.30754	0.036011	‖	0.99981

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.02	0.00741	5.3E-7	216E-19	‖	0.02741
Partially Meets the Standards	0.06217	0.4003	0.08029	6.96E-7	‖	0.54276
Meets the Standards	0.00001	0.05556	0.34672	0.00419	‖	0.40647
Exceeds the Standards	352E-18	5.62E-7	0.0135	0.00987	‖	0.02337
	0.08218	0.46326	0.44051	0.01405	‖	1.00000

Accuracy	Cut #1	Cut #2	Cut #3
0.77688	0.93041	0.86414	0.98232

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.031124	0.04889	0.00036	0.000000	‖	0.08038
Partially Meets the Standards	0.048889	0.44171	0.08531	0.000117	‖	0.57604
Meets the Standards	0.000365	0.08531	0.20789	0.013985	‖	0.30756
Exceeds the Standards	0.000000	0.00012	0.01398	0.021912	‖	0.03602
	0.080378	0.57603	0.30755	0.036013	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.031822	0.03931	0.00052	0.000000	‖	0.07167
Partially Meets the Standards	0.049980	0.35522	0.12218	0.000046	‖	0.52746
Meets the Standards	0.000373	0.06860	0.29773	0.005456	‖	0.37220
Exceeds the Standards	0.000000	0.00009	0.02003	0.008549	‖	0.02868
	0.082175	0.46324	0.44046	0.014050	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.69337	0.90980	0.80817	0.97437	‖	0.47621

Accuracy and Consistency of Classifications

Grade 8 Reading

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.039040	0.01314	0.00000	0.000000	II	0.05219
Partially Meets the Standards	0.057594	0.48676	0.04315	0.000001	II	0.58752
Meets the Standards	0.000008	0.07892	0.23679	0.010330	II	0.32605
Exceeds the Standards	0.000000	0.00000	0.00936	0.024822	II	0.03418
	0.096642	0.57882	0.28929	0.035154	II	0.99994

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.03247	0.01044	2.29E-7	694E-20	II	0.04291
Partially Meets the Standards	0.0479	0.38665	0.06622	4.94E-7	II	0.50077
Meets the Standards	6.44E-6	0.06269	0.36333	0.00468	II	0.43071
Exceeds the Standards	628E-19	6.48E-7	0.01436	0.01126	II	0.02561
	0.08037	0.45978	0.4439	0.01594	II	1.00000

Accuracy	Cut #1	Cut #2	Cut #3
0.79371	0.94166	0.87109	0.98096

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.047180	0.04926	0.00020	0.000000	II	0.09665
Partially Meets the Standards	0.049263	0.44958	0.07990	0.000120	II	0.57889
Meets the Standards	0.000198	0.07990	0.19559	0.013607	II	0.28931
Exceeds the Standards	0.000000	0.00012	0.01361	0.021427	II	0.03516
	0.096641	0.57886	0.28929	0.035155	II	1.000000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.039230	0.03912	0.00030	0.000000	II	0.07867
Partially Meets the Standards	0.040962	0.35706	0.12259	0.000055	II	0.52071
Meets the Standards	0.000165	0.06345	0.30011	0.006170	II	0.36992
Exceeds the Standards	0.000000	0.00010	0.02088	0.009716	II	0.03069
	0.080357	0.45972	0.44388	0.015941	II	1.00000

Consistency	Cut #1	Cut #2	Cut #3	II	kappa
0.70618	0.91943	0.81333	0.97280		0.50163

Accuracy and Consistency of Classifications

Grade 11 Reading

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.017342	0.00678	0.00000	0.000000	‖	0.02412
Partially Meets the Standards	0.044579	0.45819	0.05428	0.000006	‖	0.55701
Meets the Standards	0.000029	0.09744	0.26733	0.013861	‖	0.37866
Exceeds the Standards	0.000000	0.00000	0.01060	0.029488	‖	0.04009
	0.061949	0.56242	0.33221	0.043355	‖	0.99988

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.01842	0.00553	5.02E-7	109E-17	‖	0.02395
Partially Meets the Standards	0.04735	0.37357	0.07255	4.55E-6	‖	0.49348
Meets the Standards	0.00003	0.07945	0.35732	0.01011	‖	0.4469
Exceeds the Standards	868E-17	2.67E-6	0.01417	0.0215	‖	0.03567
	0.0658	0.45855	0.44403	0.03161	‖	1.00000

Accuracy	Cut #1	Cut #2	Cut #3
0.77081	0.94709	0.84797	0.97572

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.024414	0.03711	0.00043	0.000000	‖	0.06196
Partially Meets the Standards	0.037109	0.42621	0.09875	0.000326	‖	0.56244
Meets the Standards	0.000426	0.09875	0.21597	0.017052	‖	0.33225
Exceeds the Standards	0.000000	0.00033	0.01705	0.025974	‖	0.04336
	0.061950	0.56240	0.33221	0.043352	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.025928	0.03026	0.00057	0.000000	‖	0.05676
Partially Meets the Standards	0.039413	0.34747	0.13199	0.000238	‖	0.51916
Meets the Standards	0.000452	0.08052	0.28864	0.012434	‖	0.38209
Exceeds the Standards	0.000000	0.00027	0.02279	0.018940	‖	0.04200
	0.065794	0.45852	0.44398	0.031612	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.68104	0.92930	0.78595	0.96427		0.45683

Accuracy and Consistency of Classifications

Grade 4 Writing

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.12152	0.02495	0.00000	1.6627E-35	II	0.14648
Partially Meets the Standards	0.04937	0.63904	0.02987	5.7385E-12	II	0.71838
Meets the Standards	0.00000	0.03584	0.09924	.000015318	II	0.13510
Exceeds the Standards	0.00000	0.00000	0.00000	0	II	0.00000
	0.17089	0.69982	0.12911	.000015318	II	0.99997

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.08517	0.02541	721E-14	211E-36	II	0.11058
Partially Meets the Standards	0.0346	0.65082	0.03871	73E-12	II	0.72412
Meets the Standards	141E-12	0.03649	0.12861	0.00019	II	0.1653
Exceeds the Standards	0	0	0	0	II	0
	0.11977	0.71272	0.16732	0.00019	II	1

Accuracy	Cut #1	Cut #2	Cut #3
0.86460	0.94000	0.92480	0.99981

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.11890	0.05199	0.00000	6.9626E-19	II	0.17091
Partially Meets the Standards	0.05199	0.60315	0.04474	.000000028	II	0.69995
Meets the Standards	0.00000	0.04474	0.08437	.000015175	II	0.12913
Exceeds the Standards	0.00000	0.00000	0.00002	.000000118	II	0.00002
	0.17089	0.69988	0.12912	.000015321	II	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.08333	0.05294	0.00000	8.8549E-18	II	0.13628
Partially Meets the Standards	0.03644	0.61414	0.05797	.000000355	II	0.70863
Meets the Standards	0.00000	0.04556	0.10931	.000192970	II	0.15507
Exceeds the Standards	0.00000	0.00000	0.00002	.000001505	II	0.00002
	0.11977	0.71263	0.16730	.000194830	II	1.00000

Consistency	Cut #1	Cut #2	Cut #3	II	kappa
0.80687	0.91061	0.89647	0.99979		0.57337

Accuracy and Consistency of Classifications

Grade 8 Writing

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.11128	0.02127	0.00000	2.6217E-32	‖	0.13257
Partially Meets the Standards	0.04786	0.55249	0.03574	5.0676E-11	‖	0.63611
Meets the Standards	0.00000	0.04221	0.18710	.001176119	‖	0.23050
Exceeds the Standards	0.00000	0.00000	0.00043	.000324130	‖	0.00075
	0.15914	0.61597	0.22327	.001500249	‖	0.99993

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.07622	0.02073	75E-12	127E-34	‖	0.09695
Partially Meets the Standards	0.03278	0.5384	0.04642	245E-13	‖	0.6176
Meets the Standards	1.26E-9	0.04113	0.24303	0.00057	‖	0.28473
Exceeds the Standards	131E-21	426E-14	0.00056	0.00016	‖	0.00072
	0.109	0.60027	0.29001	0.00073	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.85781	0.94649	0.91245	0.99887

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.11113	0.04800	0.00000	4.6736E-17	‖	0.15916
Partially Meets the Standards	0.04800	0.51404	0.05392	.000000218	‖	0.61606
Meets the Standards	0.00000	0.05392	0.16815	.001210451	‖	0.22328
Exceeds the Standards	0.00000	0.00000	0.00121	.000289619	‖	0.00150
	0.15914	0.61596	0.22328	.001500288	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.07611	0.04678	0.00000	2.2582E-17	‖	0.12290
Partially Meets the Standards	0.03288	0.50085	0.07002	.000000105	‖	0.60386
Meets the Standards	0.00000	0.05253	0.21838	.000584960	‖	0.27153
Exceeds the Standards	0.00000	0.00000	0.00157	.000139952	‖	0.00171
	0.10899	0.60016	0.28998	.000725017	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.79562	0.92034	0.87743	0.99784	‖	0.62524

Accuracy and Consistency of Classifications

Grade 11 Writing

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.10103	0.02047	0.00000	0.000000	‖	0.12151
Partially Meets the Standards	0.03892	0.49805	0.03551	0.000000	‖	0.57239
Meets the Standards	0.00000	0.05963	0.22705	0.004705	‖	0.29138
Exceeds the Standards	-0.00000	0.00000	0.00488	0.009706	‖	0.01459
	0.13995	0.57815	0.26744	0.014412	‖	0.99987

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.06676	0.01967	2.18E-9	697E-27	‖	0.08643
Partially Meets the Standards	0.02572	0.47843	0.04299	1.2E-8	‖	0.54713
Meets the Standards	3.92E-8	0.05728	0.27485	0.00927	‖	0.3414
Exceeds the Standards	-78E-20	4.07E-9	0.00591	0.01912	‖	0.02504
	0.09248	0.55538	0.32375	0.02839	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.83916	0.95461	0.89973	0.98482

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.09795	0.04198	0.00001	0.000000	‖	0.13996
Partially Meets the Standards	0.04198	0.47247	0.06367	0.000006	‖	0.57818
Meets the Standards	0.00001	0.06367	0.19724	0.006513	‖	0.26745
Exceeds the Standards	0.00000	0.00001	0.00651	0.007895	‖	0.01441
	0.13994	0.57814	0.26743	0.014413	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.064728	0.04033	0.00001	0.000000	‖	0.10507
Partially Meets the Standards	0.027740	0.45386	0.07707	0.000011	‖	0.55871
Meets the Standards	0.000008	0.06116	0.23874	0.012829	‖	0.31277
Exceeds the Standards	0.000000	0.00001	0.00788	0.015553	‖	0.02344
	0.092476	0.55536	0.32371	0.028392	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.77292	0.93190	0.86170	0.97927	‖	0.60716

Accuracy and Consistency of Classifications

Grade 4 Mathematics

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	Marginal
Does Not Meet the Standards	0.27991	0.04060	0.00001	1.1126E-13	0.32050
Partially Meets the Standards	0.07043	0.36737	0.04371	.000008766	0.48157
Meets the Standards	0.00003	0.04543	0.14099	.006793022	0.19327
Exceeds the Standards	0.00000	0.00000	0.00213	.002448559	0.00458
	0.35037	0.45340	0.18684	.009250347	0.99991

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	Marginal
Does Not Meet the Standards	0.23006	0.04328	8.23E-6	207E-15	0.27335
Partially Meets the Standards	0.0579	0.39165	0.04947	0.00002	0.49903
Meets the Standards	0.00003	0.04844	0.15959	0.01261	0.22066
Exceeds the Standards	348E-16	1.14E-6	0.00241	0.00454	0.00695
	0.28799	0.48337	0.21148	0.01717	1

Accuracy	Cut #1	Cut #2	Cut #3
0.78585	0.89879	0.90205	0.98496

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	Marginal
Does Not Meet the Standards	0.27325	0.07642	0.00071	.000000055	0.35042
Partially Meets the Standards	0.07642	0.31598	0.06081	.000209153	0.45346
Meets the Standards	0.00071	0.06081	0.11880	.006537437	0.18687
Exceeds the Standards	0.00000	0.00021	0.00654	.002503872	0.00925
	0.35038	0.45341	0.18686	.009250517	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	Marginal
Does Not Meet the Standards	0.22458	0.08147	0.00081	0.000000	0.30686
Partially Meets the Standards	0.06281	0.33685	0.06880	0.000388	0.46887
Meets the Standards	0.00059	0.06480	0.13443	0.012135	0.21199
Exceeds the Standards	0.00000	0.00022	0.00740	0.004647	0.01227
	0.28797	0.48335	0.21144	0.017170	1.00000

Consistency	Cut #1	Cut #2	Cut #3	kappa
0.70055	0.85433	0.86436	0.97986	0.53207

Accuracy and Consistency of Classifications

Grade 8 Mathematics

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.38568	0.03958	0.00001	1.0113E-13	‖	0.42529
Partially Meets the Standards	0.05946	0.29932	0.03554	.000007236	‖	0.39435
Meets the Standards	0.00003	0.04049	0.13373	.006021500	‖	0.18027
Exceeds the Standards	0.00000	0.00000	0.00000	0	‖	0.00000
	0.44517	0.37939	0.16927	.006028736	‖	0.99991

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.3479	0.04035	7.96E-6	18E-14	‖	0.38826
Partially Meets the Standards	0.05364	0.30509	0.0422	0.00001	‖	0.40094
Meets the Standards	0.00002	0.04127	0.15881	0.01069	‖	0.2108
Exceeds the Standards	0	0	0	0	‖	0
	0.40156	0.38671	0.20102	0.01071	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.81181	0.90599	0.91648	0.98929

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.37622	0.06841	0.00058	.000000047	‖	0.44523
Partially Meets the Standards	0.06841	0.25891	0.05189	.000204325	‖	0.37944
Meets the Standards	0.00058	0.05189	0.11183	.004991531	‖	0.16930
Exceeds the Standards	0.00000	0.00020	0.00499	.000831962	‖	0.00603
	0.44520	0.37941	0.16929	.006027865	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.33929	0.06972	0.00068	0.000000	‖	0.40975
Partially Meets the Standards	0.06170	0.26385	0.06161	0.000363	‖	0.38757
Meets the Standards	0.00052	0.05288	0.13278	0.008865	‖	0.19507
Exceeds the Standards	0.00000	0.00021	0.00593	0.001477	‖	0.00761
	0.40151	0.38666	0.20100	0.010706	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.73752	0.86737	0.88373	0.98463	‖	0.59386

Accuracy and Consistency of Classifications

Grade 11 Mathematics

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.33649	0.06509	0.00002	3.0748E-13	‖	0.40161
Partially Meets the Standards	0.05598	0.33484	0.04250	.000010071	‖	0.43335
Meets the Standards	0.00002	0.03690	0.12398	.004092216	‖	0.16498
Exceeds the Standards	0.00000	0.00000	0.00000	0	‖	0.00000
	0.39249	0.43683	0.16650	.004102288	‖	0.99994

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.33664	0.05995	0.00002	774E-15	‖	0.39661
Partially Meets the Standards	0.05601	0.30841	0.04969	0.00003	‖	0.41414
Meets the Standards	0.00002	0.03398	0.14494	0.01031	‖	0.18925
Exceeds the Standards	0	0	0	0	‖	0
	0.39267	0.40235	0.19465	0.01033	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.78999	0.88400	0.91627	0.98967

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.30768	0.08420	0.00061	.000000045	‖	0.39252
Partially Meets the Standards	0.08420	0.29761	0.05485	.000193447	‖	0.43687
Meets the Standards	0.00061	0.05485	0.10751	.003525257	‖	0.16651
Exceeds the Standards	0.00000	0.00019	0.00353	.000383496	‖	0.00410
	0.39249	0.43685	0.16650	.004102246	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.30780	0.07755	0.00071	0.000000	‖	0.38609
Partially Meets the Standards	0.08423	0.27405	0.06412	0.000487	‖	0.42295
Meets the Standards	0.00061	0.05051	0.12567	0.008879	‖	0.18570
Exceeds the Standards	0.00000	0.00018	0.00412	0.000966	‖	0.00527
	0.39264	0.40229	0.19462	0.010332	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.70857	0.83688	0.88337	0.98633	‖	0.54607

Accuracy and Consistency of Classifications

Grade 4 Science and Technology

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.31641	0.04951	0.000000	4.1892E-22	‖	0.36597
Partially Meets the Standards	0.10890	0.47937	0.019348	.000000010	‖	0.60767
Meets the Standards	0.00000	0.00905	0.017296	.000007765	‖	0.02636
Exceeds the Standards	0.00000	0.00000	0.000000	0	‖	0.00000
	0.42531	0.53793	0.036644	.000007776	‖	0.99999

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.2367	0.0602	3.99E-8	608E-22	‖	0.2969
Partially Meets the Standards	0.08146	0.58292	0.01404	1.52E-6	‖	0.67841
Meets the Standards	1.04E-7	0.01101	0.01255	0.00113	‖	0.02468
Exceeds the Standards	0	0	0	0	‖	0
	0.31816	0.65413	0.02659	0.00113	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.83216	0.85834	0.97496	0.99887

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.31946	0.10580	0.000061	1.565E-12	‖	0.42537
Partially Meets the Standards	0.10580	0.41156	0.020580	.000000353	‖	0.53798
Meets the Standards	0.00006	0.02058	0.015995	.000007339	‖	0.03665
Exceeds the Standards	0.00000	0.00000	0.000007	.000000085	‖	0.00001
	0.42532	0.53795	0.036644	.000007777	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.23895	0.12863	0.000044	2.2712E-10	‖	0.36767
Partially Meets the Standards	0.07913	0.50037	0.014931	.000051230	‖	0.59457
Meets the Standards	0.00005	0.02502	0.011604	.001065016	‖	0.03774
Exceeds the Standards	0.00000	0.00000	0.000005	.000012301	‖	0.00002
	0.31813	0.65402	0.026585	.001128547	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3		kappa
0.75103	0.79211	0.95990	0.99888	‖	0.49508

Accuracy and Consistency of Classifications

Grade 8 Science and Technology

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.29279	0.04175	0.00000	6.3213E-14	‖	0.33453
Partially Meets the Standards	0.07658	0.40674	0.03568	.000007510	‖	0.51904
Meets the Standards	0.00003	0.04722	0.09628	.002633572	‖	0.14615
Exceeds the Standards	0.00000	0.00000	0.00012	.000096351	‖	0.00021
	0.36940	0.49570	0.13208	.002737433	‖	0.99994

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.24884	0.04275	4.34E-6	141E-15	‖	0.2916
Partially Meets the Standards	0.0651	0.4165	0.04655	0.00002	‖	0.52817
Meets the Standards	0.00002	0.04835	0.12563	0.00587	‖	0.17987
Exceeds the Standards	175E-17	1.28E-7	0.00015	0.00021	‖	0.00037
	0.31397	0.5076	0.17233	0.0061	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.79119	0.89213	0.90506	0.99396

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.28735	0.08154	0.00051	.000000041	‖	0.36942
Partially Meets the Standards	0.08154	0.35950	0.05449	.000170350	‖	0.49575
Meets the Standards	0.00051	0.05449	0.07486	.002219677	‖	0.13209
Exceeds the Standards	0.00000	0.00017	0.00222	.000347733	‖	0.00274
	0.36940	0.49570	0.13208	.002737801	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.24420	0.08350	0.00066	.000000091	‖	0.32839
Partially Meets the Standards	0.06931	0.36810	0.07109	.000379801	‖	0.50891
Meets the Standards	0.00043	0.05579	0.09767	.004947662	‖	0.15885
Exceeds the Standards	0.00000	0.00017	0.00290	.000775099	‖	0.00385
	0.31394	0.50757	0.17232	.006102653	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.71080	0.84609	0.87146	0.99160	‖	0.52681

Accuracy and Consistency of Classifications

Grade 11 Science and Technology

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.24478	0.05182	0.000000	7.7358E-20	‖	0.29657
Partially Meets the Standards	0.08673	0.51807	0.022991	.000000080	‖	0.62769
Meets the Standards	0.00000	0.02252	0.050964	.000740170	‖	0.07422
Exceeds the Standards	0.00000	0.00000	0.000455	.000888944	‖	0.00134
	0.33151	0.59240	0.074410	.001629194	‖	0.99982

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.2081	0.0542	1.21E-7	107E-21	‖	0.2623
Partially Meets the Standards	0.07373	0.5418	0.02978	1.11E-7	‖	0.64531
Meets the Standards	3.77E-7	0.02355	0.066	0.00102	‖	0.09058
Exceeds the Standards	167E-21	8.1E-9	0.00059	0.00123	‖	0.00182
	0.28183	0.61955	0.09637	0.00225	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.81713	0.87207	0.94667	0.99839

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.23587	0.09560	0.000065	2.0197E-11	‖	0.33154
Partially Meets the Standards	0.09560	0.46552	0.031242	.000006787	‖	0.59241
Meets the Standards	0.00006	0.03124	0.042236	.000864744	‖	0.07441
Exceeds the Standards	0.00000	0.00001	0.00865	.000757694	‖	0.00163
	0.33153	0.59236	0.074408	.001629225	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.20050	0.09998	0.000084	2.7928E-11	‖	0.30057
Partially Meets the Standards	0.08125	0.48688	0.040459	.000009384	‖	0.60863
Meets the Standards	0.00005	0.03267	0.054695	.001195669	‖	0.08863
Exceeds the Standards	0.00000	0.00001	0.001120	.001047611	‖	0.00217
	0.28181	0.61953	0.096357	.002252664	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.74315	0.81862	0.92671	0.99767	‖	0.51508

Accuracy and Consistency of Classifications

Grade 4 Social Studies

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.11223	0.03241	0.00000	0.000000	‖	0.14465
Partially Meets the Standards	0.05723	0.46997	0.06595	0.000009	‖	0.59314
Meets the Standards	0.00002	0.05781	0.18848	0.007432	‖	0.25372
Exceeds the Standards	0.00000	0.00000	0.00317	0.005207	‖	0.00837
	0.16948	0.56019	0.25760	0.012648	‖	0.99989

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.09921	0.032	5.42E-6	22E-15	‖	0.13122
Partially Meets the Standards	0.05059	0.46403	0.07273	8.74E-6	‖	0.58736
Meets the Standards	0.00002	0.05708	0.20786	0.00763	‖	0.27258
Exceeds the Standards	529E-17	7.48E-7	0.00349	0.00535	‖	0.00884
	0.14982	0.5531	0.28408	0.01299	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.77645	0.91739	0.87016	0.98886

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.10599	0.06287	0.00062	0.000000	‖	0.16949
Partially Meets the Standards	0.06287	0.41199	0.08511	0.000189	‖	0.56023
Meets the Standards	0.00062	0.08511	0.16412	0.007747	‖	0.25762
Exceeds the Standards	0.00000	0.00019	0.00775	0.004712	‖	0.01265
	0.16948	0.56016	0.25761	0.012647	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.09369	0.06207	0.00069	0.000000	‖	0.15646
Partially Meets the Standards	0.05558	0.40680	0.09386	0.000194	‖	0.55644
Meets the Standards	0.00055	0.08403	0.18097	0.007957	‖	0.27354
Exceeds the Standards	0.00000	0.00019	0.00854	0.004840	‖	0.01357
	0.14982	0.55309	0.28405	0.012991	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3		kappa
0.68632	0.88110	0.82049	0.98312	‖	0.46916

Accuracy and Consistency of Classifications

Grade 8 Social Studies

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.19495	0.02448	0.00000	0.000000	‖	0.21945
Partially Meets the Standards	0.08162	0.40198	0.04347	0.000005	‖	0.52710
Meets the Standards	0.00004	0.06009	0.06009	0.007908	‖	0.24323
Exceeds the Standards	0.00000	0.00000	0.00000	0.006212	‖	0.01019
	0.27660	0.48655	0.48655	0.014125	‖	0.99997

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.14934	0.02513	2.3E-6	817E-17	‖	0.17448
Partially Meets the Standards	0.06253	0.41264	0.05293	5.76E-6	‖	0.52811
Meets the Standards	0.00003	0.06168	0.21331	0.00983	‖	0.28485
Exceeds the Standards	114E-16	1.17E-6	0.00484	0.00772	‖	0.01256
	0.2119	0.49946	0.27109	0.01755	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.78302	0.91231	0.88535	0.98532

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.20551	0.07047	0.00063	0.000000	‖	0.27663
Partially Meets the Standards	0.07047	0.34668	0.06921	0.000190	‖	0.48657
Meets the Standards	0.00063	0.06921	0.14429	0.008509	‖	0.22267
Exceeds the Standards	0.00000	0.00019	0.00851	0.005426	‖	0.01413
	0.27660	0.48655	0.22264	0.014125	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.15741	0.07233	0.00077	0.000000	‖	0.23054
Partially Meets the Standards	0.05398	0.35583	0.08426	0.000236	‖	0.49436
Meets the Standards	0.00048	0.07104	0.17566	0.010572	‖	0.25780
Exceeds the Standards	0.00000	0.00019	0.01036	0.006743	‖	0.01730
	0.21187	0.49940	0.27105	0.017551	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3		kappa
0.69573	0.87242	0.84299	0.97864	‖	0.52011

Accuracy and Consistency of Classifications

Grade 11 Social Studies

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.29401	0.03602	0.00005	0.000000	II	0.33008
Partially Meets the Standards	0.08977	0.31299	0.04910	0.000092	II	0.45197
Meets the Standards	0.00034	0.05826	0.14655	0.011578	II	0.21671
Exceeds the Standards	0.00000	0.00000	0.00064	0.000535	II	0.00118
	0.38412	0.40727	0.19634	0.012205	II	0.99993

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.24079	0.03791	0.00006	209E-12	II	0.27877
Partially Meets the Standards	0.07352	0.32951	0.06026	0.00012	II	0.46341
Meets the Standards	0.00028	0.06133	0.17985	0.01488	II	0.25634
Exceeds the Standards	372E-13	4.01E-6	0.00078	0.00069	II	0.00147
	0.3146	0.42876	0.24096	0.01569	II	1

Accuracy	Cut #1	Cut #2	Cut #3
0.75084	0.88822	0.87794	0.98422

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.29877	0.08263	0.00271	0.000004	II	0.38413
Partially Meets the Standards	0.08263	0.25372	0.06990	0.001018	II	0.40732
Meets the Standards	0.00271	0.06990	0.11467	0.009045	II	0.19634
Exceeds the Standards	0.00000	0.00102	0.00904	0.002138	II	0.01221
	0.38411	0.40727	0.19632	0.012204	II	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.24469	0.08698	0.00332	0.000005	II	0.33501
Partially Meets the Standards	0.06767	0.26709	0.08578	0.001308	II	0.42189
Meets the Standards	0.00222	0.07358	0.14072	0.011625	II	0.22817
Exceeds the Standards	0.00000	0.00107	0.01110	0.002747	II	0.01492
	0.31458	0.42871	0.24093	0.015686	II	1.00000

Consistency	Cut #1	Cut #2	Cut #3	II	kappa
0.65530	0.83978	0.83270	0.97488	II	0.47654

Accuracy and Consistency of Classifications

Grade 4 Visual and Performing Arts

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.19098	0.05713	0.00727	0.000027	II	0.25537
Partially Meets the Standards	0.16425	0.23752	0.14764	0.006207	II	0.55554
Meets the Standards	0.00681	0.04163	0.11327	0.025360	II	0.18707
Exceeds the Standards	0.00000	0.00004	0.00084	0.000949	II	0.00183
	0.36203	0.33632	0.26901	0.032544	II	0.99981

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.14342	0.06909	0.00755	0.00004	II	0.2201
Partially Meets the Standards	0.12334	0.28725	0.15332	0.00802	II	0.57193
Meets the Standards	0.00511	0.05035	0.11762	0.03274	II	0.20583
Exceeds the Standards	9.19E-7	0.00005	0.00087	0.00123	II	0.00214
	0.27188	0.40674	0.27936	0.04202	II	1

Accuracy	Cut #1	Cut #2	Cut #3
0.54951	0.79487	0.77557	0.95829

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.20609	0.11212	0.04219	0.001630	II	0.36206
Partially Meets the Standards	0.11212	0.12766	0.08965	0.006894	II	0.33635
Meets the Standards	0.04219	0.08965	0.11870	0.018497	II	0.26905
Exceeds the Standards	0.00163	0.00689	0.01850	0.005519	II	0.03255
	0.36203	0.33632	0.26903	0.032541	II	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.15475	0.13559	0.04381	0.002105	II	0.33628
Partially Meets the Standards	0.08420	0.15439	0.09308	0.008902	II	0.34058
Meets the Standards	0.03168	010840	0.12325	0.023884	II	0.28724
Exceeds the Standards	0.00122	0.00834	0.01921	0.007126	II	0.03590
	0.27185	0.40671	0.27934	0.042017	II	1.00000

Consistency	Cut #1	Cut #2	Cut #3	II	kappa
0.43954	0.70137	0.70244	0.93633	II	0.18572

Accuracy and Consistency of Classifications

Grade 8 Visual and Performing Arts

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.27472	0.05630	0.00507	0.000111	II	0.33624
Partially Meets the Standards	0.13425	0.20200	0.08522	0.009111	II	0.43060
Meets the Standards	0.00969	0.06380	0.10269	0.040459	II	0.21664
Exceeds the Standards	0.00002	0.00075	0.00608	0.009624	II	0.01648
	0.41867	0.32284	0.19907	0.059306	II	0.99997

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.22234	0.061	0.00669	0.00009	II	0.29012
Partially Meets the Standards	0.10864	0.21889	0.11238	0.0075	II	0.44741
Meets the Standards	0.00784	0.06914	0.13541	0.0333	II	0.24569
Exceeds the Standards	0.00002	0.00081	0.00802	0.00792	II	0.01677
	0..33883	0.34985	0.26251	0.04882	II	1

Accuracy	Cut #1	Cut #2	Cut #3
0.58456	0.81572	0.79553	0.95025

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.28125	0.10553	0.02878	0.003149	II	0.41874
Partially Meets the Standards	0.10553	0.13074	0.07205	0.014502	II	0.32287
Meets the Standards	0.02878	0.07205	0.07245	0.025764	II	0.19908
Exceeds the Standards	0.00315	0.01450	0.02576	0.015892	II	0.05931
	0.41871	0.32282	0.19905	0.059307	II	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	II	Marginal
Does Not Meet the Standards	0.22757	0.11435	0.03795	0.002592	II	0.38250
Partially Meets the Standards	0.08539	0.14166	0.09502	0.011936	II	0.33404
Meets the Standards	0.02329	0.07808	0.09554	0.021206	II	0.21814
Exceeds the Standards	0.00255	0.01571	0.03397	0.013081	II	0.06532
	0.33879	0.34980	0.26247	0.048814	II	1.00000

Consistency	Cut #1	Cut #2	Cut #3	II	kappa
0.47791	0.73386	0.73285	0.91203	II	0.24671

Accuracy and Consistency of Classifications

Grade 11 Visual and Performing Arts

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.46533	0.12317	0.01184	0.000028	II	0.60034
Partially Meets the Standards	0.05456	0.07590	0.04706	0.002052	II	0.17957
Meets the Standards	0.00577	0.04258	0.13925	0.032394	II	0.22000
Exceeds the Standards	0.00000	0.00000	0.00000	0.000000	II	0.00000
	0.52566	0.24165	0.19815	0.034475	II	0.99991

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.36453	0.14996	0.0166	0.00001	II	0.5311
Partially Meets the Standards	0.04274	0.09241	0.06596	0.00097	II	0.20208
Meets the Standards	0.00452	0.05185	0.19519	0.01526	II	0.26682
Exceeds the Standards	0	0	0	0	II	0
	0.41179	0.29422	0.27776	0.01624	II	1

Accuracy	Cut #1	Cut #2	Cut #3
0.65213	0.78617	0.86009	0.98376

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.38330	0.11844	0.02288	0.001014	II	0.52569
Partially Meets the Standards	0.11844	0.06769	0.04904	0.006474	II	0.24166
Meets the Standards	0.02288	0.04904	0.10457	0.021652	II	0.19816
Exceeds the Standards	0.00101	0.00647	0.02165	0.005336	II	0.03448
	0.52563	0.24164	0.19814	0.034476	II	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.30023	0.14420	0.03207	0.000477	II	0.47704
Partially Meets the Standards	0.09277	0.08241	0.06874	0.003049	II	0.24699
Meets the Standards	0.01792	0.05971	0.14658	0.010199	II	0.23442
Exceeds the Standards	0.00079	0.00788	0.03035	0.002513	II	0.04154
	0.41172	0.29420	0.27774	0.016238	II	1.00000

Consistency	Cut #1	Cut #2	Cut #3		kappa
0.53182	0.71175	0.80935	0.94724	II	0.29608

Accuracy and Consistency of Classifications

Grade 4 Health Education

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.000000	0.00000	0.00000	0.000000	II	0.00000
Partially Meets the Standards	0.026215	0.59473	0.13361	0.000108	II	0.75464
Meets the Standards	0.000163	0.06473	0.15802	0.009138	II	0.23206
Exceeds the Standards	0.000000	0.00001	0.00409	0.009136	II	0.01323
	0.026377	0.65946	0.29572	0.018382	II	0.999993

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0	0	0	0	II	0
Partially Meets the Standards	0.02212	0.61179	0.12327	0.00016	II	0.75734
Meets the Standards	0.00014	0.06659	0.14581	0.01318	II	0.22571
Exceeds the Standards	525E-14	6.12E-6	0.00378	0.01318	II	0.01696
	0.02226	0.67838	0.27285	0.02651	II	1

Accuracy	Cut #1	Cut #2	Cut #3
0.77077	0.97774	0.80984	0.98288

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.001242	0.02183	0.00330	0.000002	II	0.02638
Partially Meets the Standards	0.021832	0.49854	0.13846	0.000637	II	0.65949
Meets the Standards	0.003303	0.13846	0.14456	0.009380	II	0.29574
Exceeds the Standards	0.000002	0.00064	0.00938	0.008364	II	0.01838
	0.026378	0.65946	0.29570	0.018382	II	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.001048	0.02246	0.00305	0.000002	II	0.02656
Partially Meets the Standards	0.018421	0.51282	0.12775	0.000918	II	0.65992
Meets the Standards	0.002786	0.14243	0.13339	0.013527	II	0.29215
Exceeds the Standards	0.000001	0.00065	0.00865	0.012060	II	0.02137
	0.022257	0.67835	0.27284	0.026508	II	1.00000

Consistency	Cut #1	Cut #2	Cut #3		kappa
0.65933	0.95328	0.72240	0.97624	II	0.27740

Accuracy and Consistency of Classifications

Grade 8 Health Education

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.000000	0.00000	0.00000	0	‖	0.00000
Partially Meets the Standards	0.026676	0.57971	0.11711	.000050783	‖	0.72351
Meets the Standards	0.000115	0.08247	0.18726	.005617142	‖	0.27545
Exceeds the Standards	0.000000	0.00000	0.00045	.000438035	‖	0.00089
	0.026791	0.66219	0.30482	.006105959	‖	0.99985

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0	0	0	0	‖	0
Partially Meets the Standards	0.0307	0.5362	0.13553	0.00003	‖	0.70246
Meets the Standards	0.00013	0.07628	0.21671	0.00362	‖	0.29674
Exceeds the Standards	134E-15	4.71E-7	0.00052	0.00028	‖	0.0008
	0.03084	0.61248	0.35275	0.00394	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.75319	0.96916	0.78803	0.99583

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.02679	0.001794	0.02293	0.00207	‖	.000000468
Partially Meets the Standards	0.66226	0.022926	0.50269	0.13611	‖	.000435174
Meets the Standards	0.30484	0.002070	0.13611	0.16183	‖	.004772186
Exceeds the Standards	0.00611	0.000000	0.00044	0.00477	‖	.000898123
	1.00000	0.026791	0.66216	0.30479	‖	.006105950

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards	‖	Marginal
Does Not Meet the Standards	0.002065	0.02120	0.00240	.000000301	‖	0.02567
Partially Meets the Standards	0.026386	0.46497	0.15750	.000280440	‖	0.64917
Meets the Standards	0.002382	0.12589	0.18729	.003075123	‖	0.31866
Exceeds the Standards	0.000001	0.00040	0.00552	.000578761	‖	0.00650
	0.030834	0.61246	0.35271	.003934625	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.65492	0.94763	0.71111	0.99072	‖	0.29457

Accuracy and Consistency of Classifications

Grade 11 Health Education

Step 4
Predicted Classification X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.000000	0.00000	0.00000	0.000000	‖	0.00000
Partially Meets the Standards	0.036797	0.59546	0.14243	0.000235	‖	0.77490
Meets the Standards	0.000276	0.06070	0.14355	0.010008	‖	0.21454
Exceeds the Standards	0.000000	0.00001	0.00359	0.006898	‖	0.01049
	0.037072	0.65617	0.28957	0.017141	‖	0.99994

Step 5
Actual Classification X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0	0	0	0	‖	0
Partially Meets the Standards	0.04546	0.55566	0.16412	0.00011	‖	0.76535
Meets the Standards	0.00034	0.05665	0.16543	0.00479	‖	0.22721
Exceeds the Standards	716E-13	9.74E-6	0.00413	0.0033	‖	0.00745
	0.0458	0.61232	0.33368	0.00821	‖	1

Accuracy	Cut #1	Cut #2	Cut #3
0.72439	0.95420	0.77877	0.99095

Step 6
X(1)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.002264	0.02970	0.00510	0.000006	‖	0.03707
Partially Meets the Standards	0.029697	0.48486	0.14056	0.000990	‖	0.65619
Meets the Standards	0.005104	0.14056	0.13431	0.009583	‖	0.28959
Exceeds the Standards	0.000006	0.00099	0.00958	0.006562	‖	0.01714
	0.037072	0.65611	0.28956	0.017141	‖	1.00000

Step 7
X(0)

tstat	Does Not Meet the Standards	Partially Meets the Standards	Meets the Standards	Exceeds the Standards		Marginal
Does Not Meet the Standards	0.002797	0.02771	0.00588	.000002761	‖	0.03640
Partially Meets the Standards	0.036682	0.45245	0.16196	.000474036	‖	0.65164
Meets the Standards	0.006305	0.13116	0.15475	.004588127	‖	0.29685
Exceeds the Standards	0.000007	0.00092	0.01104	.003141880	‖	0.01512
	0.045791	0.61226	0.33363	.008206804	‖	1.00000

Consistency	Cut #1	Cut #2	Cut #3	‖	kappa
0.61321	0.92340	0.69324	0.98296	‖	0.22665

APPENDIX B

SAMPLE REPORTS

Important Information for the Parents/Guardians of

_____

Grade 8 Assessment
March 2000 Administration

MAINE EDUCATIONAL ASSESSMENT

STATE OF MAINE
DEPARTMENT OF EDUCATION
23 State House Station
Augusta, ME 04333

September 2000

J. Duke Albanese
COMMISSIONER

Dear Parents/ Guardians:

The Legislature approved Maine's *Learning Results* in May of 1997, giving all schools standards to measure student learning. Our goal is for all students in Maine to demonstrate that they meet all standards defined in the *Learning Results.* The Maine Educational Assessment (MEA), the state test which has been administered for the past 16 years in grades 4, 8, and 11, has been rewritten so it tests the challenging subject matter specified in the *Learning Results.* Your student was in the second group to take this new test. In the past, the MEA reported individual student scores only in reading and writing (called English/Language Arts in the *Learning Results)* and in mathematics. The new MEA also includes individual results in science and technology, and in social studies.

During March of 2000, students in grades 4, 8, and 11 were tested in mathematics, science and technology, and social studies as part of the MEA. The test included multiple-choice questions, short-answer questions, and essay questions (known as "constructed response"). The report on the reverse side of this letter provides you with important information about your student's performance on this part of the MEA, along with a summary of school, district, and state results. (An earlier MEA report included results for reading and writing.) Please keep in mind that your student's score measures learning over the past 3-4 years, not just the work of the past year.

Staff at your school can provide further information about school and district results, and about your student's performance overall. The MEA is just one part of the comprehensive assessment system your school uses to measure student learning and school success. MEA results are used at the school, district, and state levels to improve teaching and learning.

Sincerely,

J. Duke Albanese
Commissioner

Information on Maine's *Learning Results*

The *Learning Results* were developed in eight content areas by thousands of Maine citizens.

The MEA was rewritten by hundreds of Maine teachers to align with the *Learning Results.*

Setting MEA performance standards based on the quality of student work was completed by hundreds of Maine teachers and citizens.

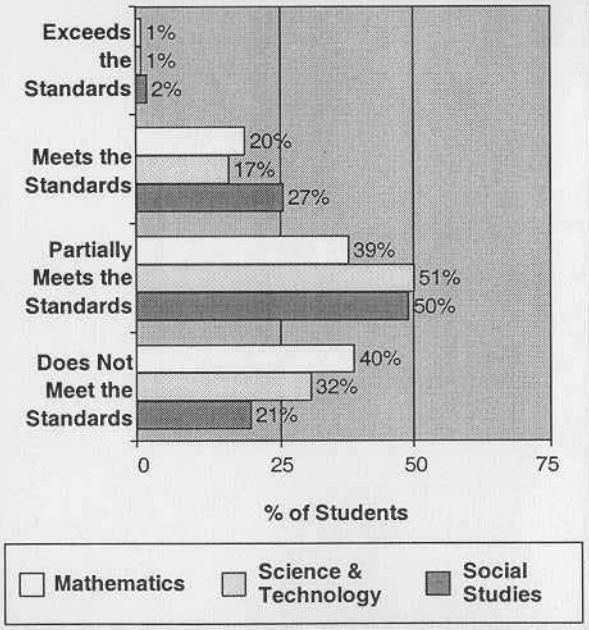
For a copy of Maine's *Learning Results* either call 287-4468 or find them on-line at http://janus.state.me.us/education/lres/homepage.htm

Performance Levels and Score Ranges

On this assessment, results are reported as four performance levels using scaled scores that range from 501 to 580. **The chart below describes the quality of student work:**

- ☐ **Exceeds the Standards (561 to 580)**
  The student's work demonstrates exemplary accomplishment of content knowledge, analysis, problem solving, and communication skills.
- ☐ **Meets the Standards (541 to 560)**
  The student's work demonstrates consistent accomplishment of content knowledge, analysis, problem solving, and communication skills.
- ☐ **Partially Meets the Standards (521 to 540)**
  The student's work demonstrates inconsistent accomplishment of content knowledge, analysis, problem solving, and communication skills.
- ☐ **Does Not Meet the Standards (501 to 520)**
  The student's work demonstrates limited command of content knowledge, analysis, problem solving, and communication skills.

Maine State MEA Summary Results for March 2000 Administration

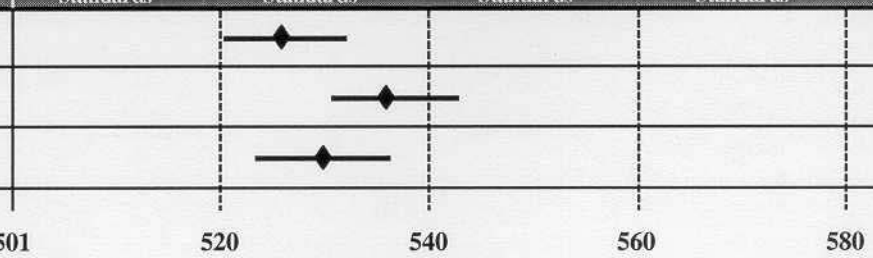


00737408

Student	Grade	School	District
[redacted]	8	[redacted]	[redacted]

Content Area	Performance Level	Score
Mathematics	Partially Meets the Standards	526
Science & Technology	Partially Meets the Standards	536
Social Studies	Partially Meets the Standards	530

This Student's Performance Levels and Scores

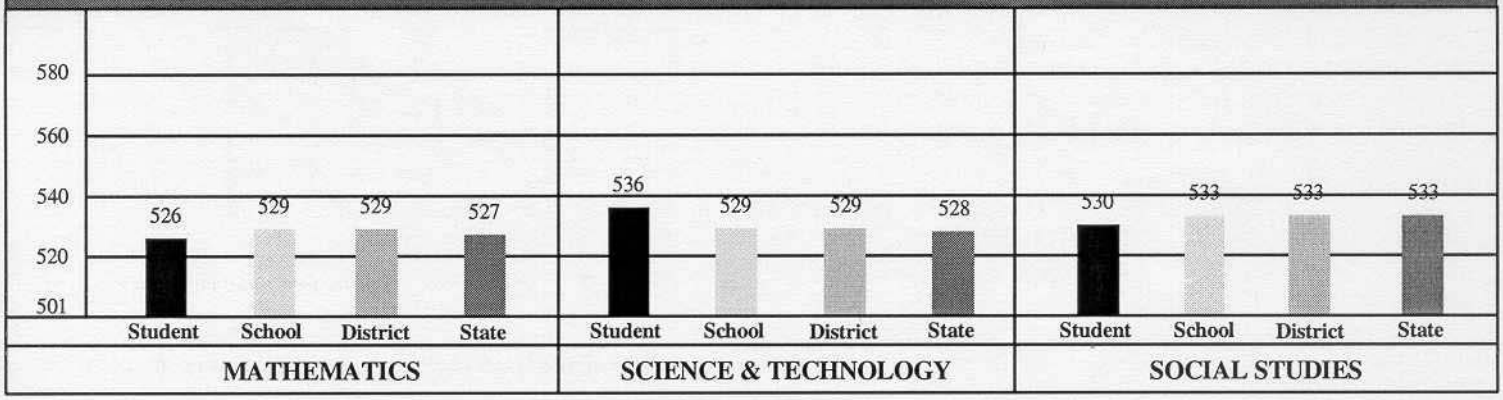


Testing Incomplete (TI): Student failed to attempt one or more sessions for that content area.

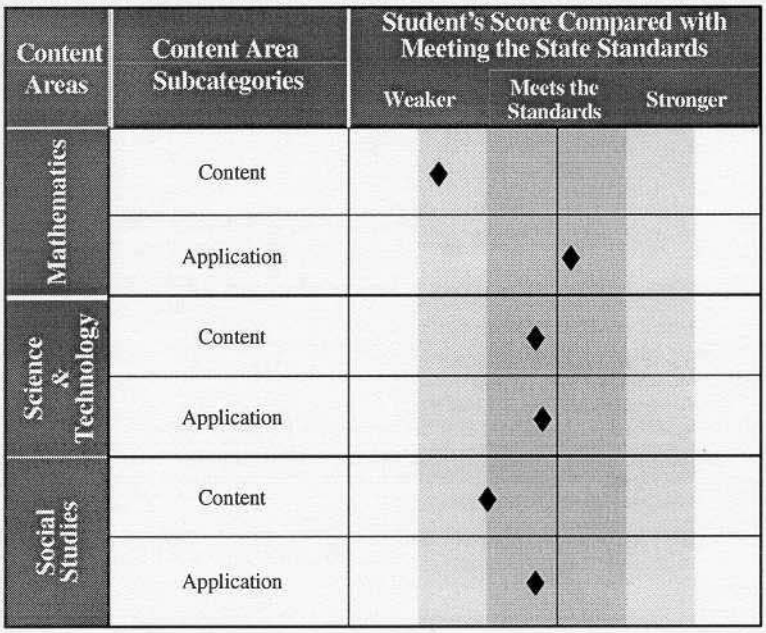
See reverse side for description of performance levels and state summary results.

The ◆ represents the student's score. The bar (——) surrounding the score represents the probable range of scores for the student if he or she was tested many times. This statistic is called the standard error of measurement.

How this Student's Performance Compared to Average Scores from School, District, and State



This Student's Performance in Content Area Subcategories



Definitions of Content Area Subcategories

Content: Refers to a student's knowledge and conceptual understanding of the content area and of the procedures necessary to acquire new learning.

Application: Refers to a student's use of knowledge and to his/her conceptual and procedural understanding for applying knowledge in the content area through reasoning, inquiry, communicating ideas, and/or solving problems.

Scores for Content and Application are derived from particular subsets of items in each content area that emphasize those types of knowledge.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Partially Meets	526
Date: 03/00	Science:	Partially Meets	536
	Social Studies:	Partially Meets	530

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Partially Meets	536
Date: 03/00	Science:	Partially Meets	532
	Social Studies:	Meets	544

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Partially Meets	534
Date: 03/00	Science:	Partially Meets	536
	Social Studies:	Partially Meets	538

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Meets	548
Date: 03/00	Science:	Partially Meets	534
	Social Studies:	Partially Meets	530

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Partially Meets	530
Date: 03/00	Science:	Partially Meets	536
	Social Studies:	Meets	546

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Partially Meets	536
Date: 03/00	Science:	Partially Meets	528
	Social Studies:	Partially Meets	538

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Partially Meets	540
Date: 03/00	Science:	Partially Meets	532
	Social Studies:	Meets	548

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Does Not Meet	514
Date: 03/00	Science:	Does Not Meet	512
	Social Studies:	Does Not Meet	516

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Meets	550
Date: 03/00	Science:	Partially Meets	540
	Social Studies:	Partially Meets	540

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

Name:
School:
District:

		Performance Levels	Scaled Scores
Grade: 8	Mathematics:	Does Not Meet	502
Date: 03/00	Science:	Does Not Meet	508
	Social Studies:	Does Not Meet	504

The MEA was revised in 1998/99 to assess Maine's *Learning Results,* required by law to be fully implemented by 2002-2003.

DEPARTMENT OF EDUCATION

1999-2000 School Year Reports

Dear School Board Members and School Personnel:

The Maine Educational Assessment (MEA) is in its second year of measuring student performance on Maine's *Learning Results,* which challenge schools and students to pursue academic standards that are among the highest in the nation. This report of student performance in Mathematics, Science and Technology, Social Studies, and Visual and Performing Arts on tests administered in March 2000 is the second of two summary reports for the school year 1999-2000. You have already received student performance results in Reading, Writing, and Health Education. This MEA results report should still be considered baseline information, as the *Learning Results* are not scheduled for full implementation until the 2003 school year.

The MEA, revised to align with Maine's *Learning Results,* is composed of selected-response (multiple choice), short answer, and complex questions, including a writing prompt, that require students to construct answers that demonstrate their knowledge and skills. Your review of the MEA questions that we have released will help you understand the revised assessment and the challenge that it presents for students and schools. The scores are reported using a numerical scale (501-580) and performance levels of Does Not Meet the Standards, Partially Meets the Standards, Meets the Standards, and Exceeds the Standards. The scale and the performance levels, established in the fall of 1999, will remain fixed for a period of at least five years to measure progress of students across the state in achieving the standards. It is important to know that more than 500 teachers and other educators from across Maine helped to develop the revised MEA and assisted in the scoring and standard-setting process. This participation has not only strengthened the redesigned MEA but has also engaged teachers from around the state in conversations about quality standards for student work.

I look forward to continuing the strong state and local partnership that has led to our current success as we work toward achieving even higher standards for all Maine students.

Sincerely,

J. Duke Albanese
Commissioner

MAINE EDUCATIONAL ASSESSMENT

Educational Assessment Report

ID:
School:
District:
Grade: 8
Test Date: March 2000

Contents of the Report

The report is divided into five main sections including a section describing the students tested and a separate section for the results in each content area.

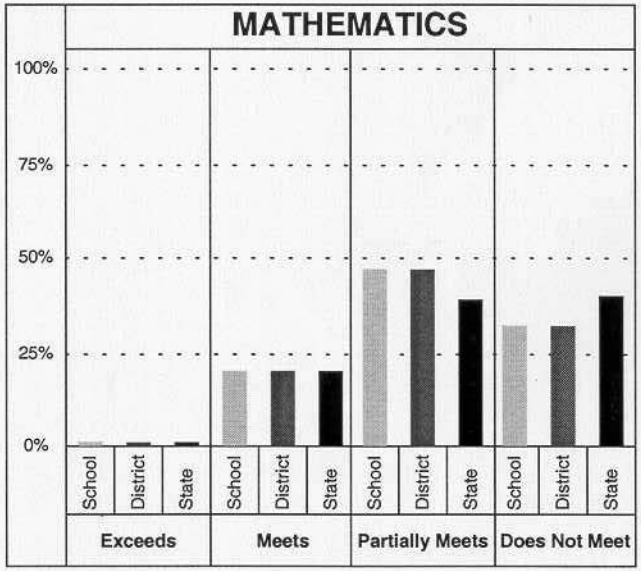
MAINE EDUCATIONAL ASSESSMENT

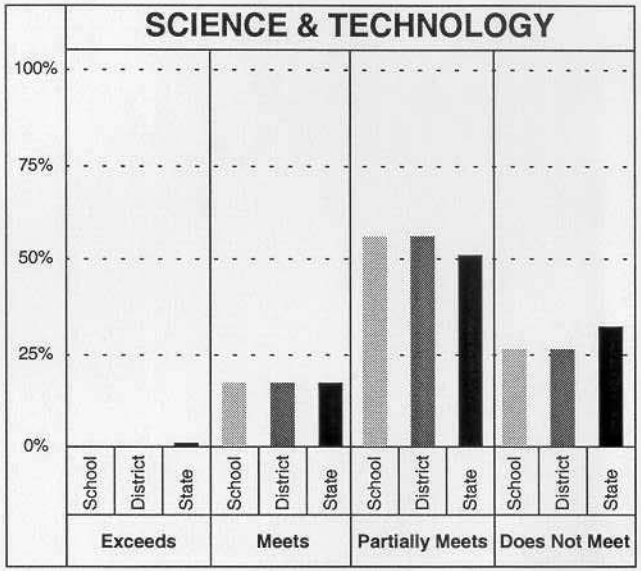
SUMMARY OF SCORES

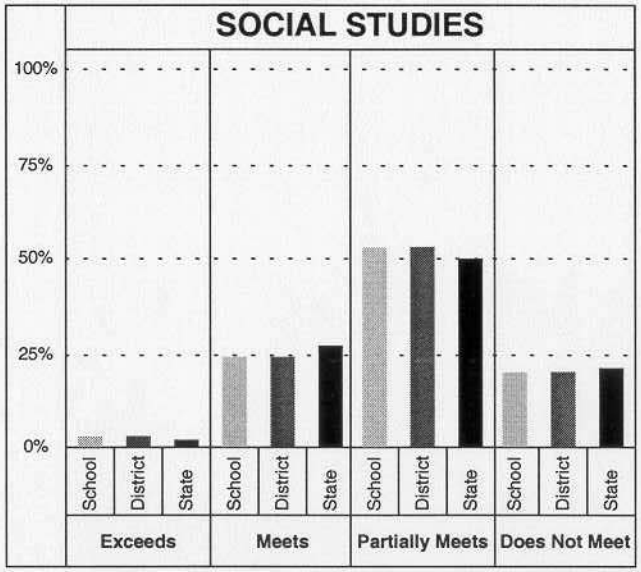
School:
District:
Grade: 8
Date: March 2000

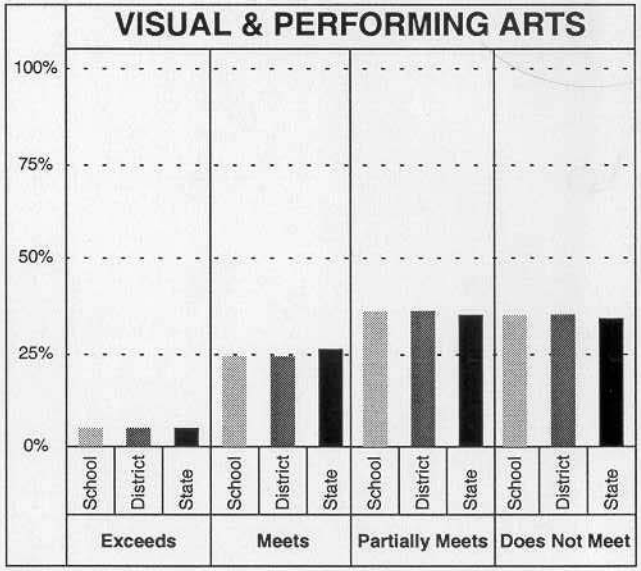
Executive Summary of School, District, and State Scores

Year	Average Performance Score		
	School	District	State
MATHEMATICS			
1998-1999	526	526	529
1999-2000	**529**	**529**	**527**
2000-2001			
Cum. Avg.	528	528	528
SCIENCE & TECHNOLOGY			
1998-1999	527	527	528
1999-2000	**529**	**529**	**528**
2000-2001			
Cum. Avg.	528	528	528
SOCIAL STUDIES			
1998-1999	533	533	532
1999-2000	**533**	**533**	**533**
2000-2001			
Cum. Avg.	533	533	532
VISUAL & PERFORMING ARTS			
1998-1999	536	536	532
1999-2000	**530**	**530**	**530**
2000-2001			
Cum. Avg.	533	533	531









MAINE EDUCATIONAL ASSESSMENT

SUMMARY OF STUDENT PARTICIPATION

School: [redacted]
District: [redacted]
Grade: 8
Date: March 2000

On this page, several statistics concerning the number of students who were enrolled in the school and those who actually were tested are reported. Schools with large percentages of non-tested students must interpret their results with caution because results for the students they tested may not be representative of their total school performance. Because sufficient time was provided for makeup testing, schools were expected to administer the full battery of tests to all students with the exception of some students with an identified disability, students with limited-English proficiency who could not meaningfully respond to the test, and students who were chronically absent from school.

Participation Category	Number			Percentage		
	State	District	School	State	District	School
Students Enrolled: number of returned test booklets	17602	262	262	100	100	100
Students Excluded from Report(s): students totally excluded from testing (took no session of the assessment) due to an identified disability	164	1	1	1	0	0
students partially excluded from testing (excluded from some but not all sessions of the assessment) due to an identified disability	24	0	0	0	0	0
students tested, but excluded from report because they receive special education and related services for more than 60% of the school day in a composite or self-contained program (categories 24 or 25 on EF-S-204)	233	7	7	1	3	3
students totally excluded from testing because of LEP, Title 1 decision or other approved reason	54	0	0	0	0	0
students partially excluded from testing because of LEP, Title 1 decision or other approved reason	4	0	0	0	0	0
others totally excluded from testing	364	11	11	2	4	4
others partially excluded from testing	335	2	2	2	1	1
Students with Identified Disability Completing All Subjects without Accommodations	424	4	4	2	2	2
Students with Identified Disability Completing All Subjects with Accommodations	1308	16	16	7	6	6
All Others Completing All Subjects	14692	221	221	83	84	84
Total Completing All Subjects	16424	241	241	93	92	92
Percentage of Students with Identified Disability Included in Reports for All Subjects				83	79	79
Percentage of All Other Students Included in Reports for All Subjects				95	94	94

MAINE EDUCATIONAL ASSESSMENT

MATHEMATICS RESULTS

School:
District:
Grade: 8
Date: March 2000

PERFORMANCE LEVELS		STUDENTS AT EACH PERFORMANCE LEVEL				
		School		District		State
		N	%	N	%	%
Exceeds the Standards - The quality of a student's work at this level of proficiency exceeds the standards of performance as identified for Maine's *Learning Results* in mathematics. The student's overall performance demonstrates exemplary knowledge of content, process, problem-solving, reasoning and communication skills. Scaled scores (561-580).	1998-1999	2	1	2	1	1
	1999-2000	**2**	**1**	**2**	**1**	**1**
	2000-2001					
	Cumulative Average	2	1	2	1	1
Meets the Standards - The quality of a student's work at this level of proficiency meets the standards of performance as identified for Maine's *Learning Results* in mathematics. The student's body of work consistently shows complete knowledge of mathematical content, process, reasoning and communication skills, and problem-solving ability. Scaled scores (541-560).	1998-1999	29	13	29	13	21
	1999-2000	**49**	**20**	**49**	**20**	**20**
	2000-2001					
	Cumulative Average	39	17	39	17	21
Partially Meets the Standards - The quality of a student's work at this level of proficiency partially meets the standards of performance as identified for Maine's *Learning Results* in mathematics. The student's body of work demonstrates a partial and/or inconsistent knowledge of mathematical content, process, reasoning and communication skills, and problem-solving ability. Scaled scores (521-540).	1998-1999	108	48	108	48	41
	1999-2000	**114**	**47**	**114**	**47**	**39**
	2000-2001					
	Cumulative Average	111	48	111	48	40
Does Not Meet the Standards - The quality of a student's work at this level of proficiency does not meet the standards of performance as identified for Maine's *Learning Results* in mathematics. The student's body of work demonstrates a limited knowledge of mathematical content, process, reasoning and communication skills, and problem-solving ability. Scaled scores (501-520).	1998-1999	88	39	88	39	37
	1999-2000	**77**	**32**	**77**	**32**	**40**
	2000-2001					
	Cumulative Average	83	36	83	36	39

Learning Results Content Standards	Number of Points Possible	Average Points Attained (Number and Percent)					
		School		District		State	
		N	%	N	%	N	%
Content	117	56.8	49	56.8	49	54.1	46
Application	89	35.5	40	35.5	40	31.6	36
Numbers and Number Sense (Standard A)	28	11.9	43	11.9	43	11.5	41
Computation (Standard B)	23	10.1	44	10.1	44	9.7	42
Data Analysis and Statistics (Standard C)	22	11.4	52	11.4	52	10.9	50
Probability (Standard D)	23	10.2	44	10.2	44	9.4	41
Geometry (Standard E)	24	9.7	40	9.7	40	9.0	38
Measurement (Standard F)	22	9.3	42	9.3	42	9.3	42
Patterns, Relations, Functions (Standard G)	26	13.2	51	13.2	51	11.6	45
Algebra Concepts (Standard H)	29	12.2	42	12.2	42	10.2	35
Discrete Mathematics (Standard I)	9	4.4	49	4.4	49	4.1	46

MAINE EDUCATIONAL ASSESSMENT

MATHEMATICS RESULTS

(CONTINUED)

School:
District:
Grade: 8
Date: March 2000

Reporting Categories	School					State				
	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
GENDER										
boy	48	0	19	46	35	51	1	21	38	40
girl	52	2	21	49	28	49	1	20	40	40
INTERNET-CONNECTED COMPUTER IN HOME										
yes	74	1	22	48	30	72	1	23	40	36
no	24	2	17	47	34	27	0	13	36	50
I don't know.						1	0	7	31	62
QUANTITY OF TV ON A SCHOOL NIGHT										
none	4	0	10	60	30	6	2	27	37	34
less than one hour	18	0	28	40	33	25	1	24	40	34
one to two hours	51	2	24	50	24	43	1	21	41	37
more than two hours	27	0	9	46	45	27	0	13	34	52
H.O.T.S.										
students in a H.O.T.S. program						2	1	9	26	64
students completed program 2 years prior						1	0	5	28	67
TITLE 1 PROGRAM										
students currently served in math						2	0	5	19	76
students previously served in math						3	0	6	26	67
new students currently served in reading						1	0	1	17	82
new students previously served in reading	7	0	0	18	82	2	0	3	19	78
MIGRANT										
students eligible, not served						0	0	20	32	48
students eligible, served, not tutored						1	0	15	24	60
students eligible, served, tutored						1	0	4	35	61
STATE-APPROVED GIFTED/ TALENTED PROGRAM										
yes						4	10	68	18	5
no	100	1	20	47	32	96	1	18	40	42
ENGLISH LANGUAGE FLUENCY										
monolingual English	96	0	17	55	28	98	1	20	39	40
bilingual/English fluent						1	2	21	31	47
bilingual/limited-English proficient						1	0	13	28	59
IDENTIFIED DISABILITY										
yes	9	0	5	29	67	11	0	2	16	82
no	91	1	22	49	29	89	1	22	41	35
OPTIONAL SCHOOL/DISTRICT QUESTION										
A										
B										
C										
D										
E										

Questionnaire Items	Sch.	State				
	% Students in Each Category	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
How well can you learn the mathematics that is taught in your school?						
very well	28	36	2	30	40	27
well	55	49	1	17	40	43
not too well	13	12	0	9	33	58
not well at all	4	4	0	7	23	70
What best describes your mathematics classes?						
The teacher talks about mathematics and I work by myself to do assignments from the book.	29	31	1	21	39	39
The teacher talks about mathematics and I work by myself to investigate and solve problems.	18	11	0	12	35	52
The teacher talks about mathematics and we work in groups to investigate and solve problems.	5	9	1	13	33	53
a combination of the options above	49	48	1	23	41	35
"I learn in school most of what I need to know to answer the MEA mathematics questions."						
strongly agree	23	31	3	35	36	26
agree	60	51	0	16	42	41
disagree	14	13	0	7	35	58
strongly disagree	3	5	0	6	26	68
My grades in mathematics depend mostly on						
tests and quizzes.	6	14	1	20	36	43
tests, quizzes, and homework.	80	63	1	21	40	38
journals and portfolios.	1	2	1	10	24	65
a combination of the options above.	13	21	1	18	39	42
"My knowledge of mathematics will be useful to me in my future work."						
strongly agree	59	61	1	24	40	35
agree	32	32	1	15	38	46
disagree	7	5	0	12	31	57
strongly disagree	2	2	0	7	25	67
What best describes the use of calculators in your mathematics classes?						
Calculators are used daily.	13	38	2	27	37	34
Calculators are used once or twice a week.	45	35	1	18	41	40
Calculators are used once or twice a month.	42	19	1	14	39	46
Calculators are never used.	0	9	0	12	36	52
What best describes the use of computers in your mathematics classes?						
Computers are used daily.	2	4	0	8	29	62
Computers are used once or twice a week.	2	7	1	15	29	54
Computers are used once or twice a month.	10	17	2	23	38	36
Computers are never used.	86	72	1	21	40	38
What best describes the mathematics class you are taking in the eighth-grade?						
basic mathematics	4	22	0	4	31	66
advanced mathematics	3	13	0	12	39	49
pre-algebra	56	41	0	12	48	40
Algebra I	36	24	4	52	31	13

MAINE EDUCATIONAL ASSESSMENT

SCIENCE & TECHNOLOGY RESULTS

School:
District:
Grade: 8
Date: March 2000

PERFORMANCE LEVELS	STUDENTS AT EACH PERFORMANCE LEVEL					
		School		District		State
		N	%	N	%	%
Exceeds the Standards - The quality of a student's work at this level of proficiency exceeds the standards of performance as identified for Maine's *Learning Results* in science and technology. The student demonstrates exemplary knowledge of content including life, physical, and earth/space sciences and scientific inquiry, reasoning, and communication skills. Scaled scores (561-580).	1998-1999	0	0	0	0	<1
	1999-2000	**1**	**0**	**1**	**0**	**1**
	2000-2001					
	Cumulative Average	1	0	1	0	1
Meets the Standards - The quality of a student's work at this level of proficiency meets the standards of performance as identified for Maine's *Learning Results* in science and technology. The student demonstrates consistent knowledge of content including life, physical, and earth/space sciences and scientific inquiry, reasoning, and communication skills. Scaled scores (541-560).	1998-1999	27	12	27	12	14
	1999-2000	**42**	**17**	**42**	**17**	**17**
	2000-2001					
	Cumulative Average	35	15	35	15	16
Partially Meets the Standards - The quality of a student's work at this level of proficiency partially meets the standards of performance as identified for Maine's *Learning Results* in science and technology. The student demonstrates partial and/or inconsistent knowledge of content including life, physical, and earth/space sciences and scientific inquiry, reasoning, and communication skills. Scaled scores (521-540).	1998-1999	133	59	133	59	56
	1999-2000	**136**	**56**	**136**	**56**	**51**
	2000-2001					
	Cumulative Average	135	58	135	58	54
Does Not Meet the Standards - The quality of a student's work at this level of proficiency does not meet the standards of performance as identified for Maine's *Learning Results* in science and technology. The student demonstrates limited knowledge of content including life, physical, and earth/space sciences and scientific inquiry, reasoning, and communication skills. Scaled scores (501-520).	1998-1999	66	29	66	29	30
	1999-2000	**64**	**26**	**64**	**26**	**32**
	2000-2001					
	Cumulative Average	65	28	65	28	31

Learning Results Content Standards	Number of Points Possible	Average Points Attained (Number and Percent)					
		School		District		State	
		N	%	N	%	N	%
Content	136	64.2	47	64.2	47	65.1	48
Classifying Life Forms (Standard A)	11	6.3	57	6.3	57	6.1	55
Ecology (Standard B)	14	6.5	46	6.5	46	7.4	53
Cells (Standard C)	18	7.0	39	7.0	39	7.4	41
Continuity and Change (Standard D)	17	7.1	42	7.1	42	7.4	44
Structure of Matter (Standard E)	18	9.0	50	9.0	50	8.0	44
The Earth (Standard F)	20	8.4	42	8.4	42	9.1	46
The Universe (Standard G)	15	6.8	45	6.8	45	6.6	44
Energy (Standard H)	10	4.2	42	4.2	42	4.6	46
Motion (Standard I)	13	8.9	68	8.9	68	8.4	65
Application	58	22.9	39	22.9	39	22.2	38
Inquiry and Problem Solving (Standard J)	19	10.8	57	10.8	57	9.5	50
Scientific Reasoning (Standard K)	18	5.9	33	5.9	33	5.9	33
Communication (Standard L)	11	3.9	35	3.9	35	4.0	36
Implications of Science & Technology (Standard M)	10	2.3	23	2.3	23	2.8	28

MAINE EDUCATIONAL ASSESSMENT

SCIENCE & TECHNOLOGY RESULTS

(CONTINUED)

School:
District:
Grade: 8
Date: March 2000

Reporting Categories	School					State				
	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
GENDER										
boy	48	0	18	57	25	51	1	19	51	29
girl	52	1	16	56	27	49	1	15	51	34
INTERNET-CONNECTED COMPUTER IN HOME										
yes	74	1	17	59	23	72	1	20	52	27
no	24	0	19	46	36	27	0	11	48	41
I don't know.						1	1	8	40	52
QUANTITY OF TV ON A SCHOOL NIGHT										
none	4	0	10	60	30	6	1	25	50	23
less than one hour	18	0	28	47	26	25	1	20	51	28
one to two hours	51	1	17	61	21	43	1	18	53	28
more than two hours	27	0	12	51	37	27	0	11	47	42
H.O.T.S.										
students in a H.O.T.S. program						2	0	5	38	57
students completed program 2 years prior						1	0	6	44	50
TITLE 1 PROGRAM										
students currently served in math						2	0	3	32	65
students previously served in math						3	0	6	41	53
new students currently served in reading						1	0	1	26	73
new students previously served in reading	7	0	6	47	47	2	0	3	31	66
MIGRANT										
students eligible, not served						0	0	17	50	33
students eligible, served, not tutored						1	0	10	36	54
students eligible, served, tutored						1	0	6	36	58
STATE-APPROVED GIFTED/ TALENTED PROGRAM										
yes						4	6	61	29	4
no	100	0	17	56	26	96	0	15	52	33
ENGLISH LANGUAGE FLUENCY										
monolingual English	96	0	16	61	23	98	1	17	51	31
bilingual/English fluent						1	0	12	49	39
bilingual/limited-English proficient						1	0	6	40	54
IDENTIFIED DISABILITY										
yes	9	0	5	50	45	11	0	3	31	66
no	91	0	19	57	24	89	1	19	53	27
OPTIONAL SCHOOL/DISTRICT QUESTION										
A										
B										
C										
D										
E										

Questionnaire Items	Sch.	State				
	% Students in Each Category	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
Which statement best describes how your science class is taught?						
We read text, answer questions, and other activities.	54	32	1	17	49	33
We use materials to design our labs and activities.	1	8	0	11	45	43
We have mostly lectures and demonstrations.	7	11	1	15	51	34
We have a balanced combination of options above.	38	49	1	19	53	27
Which statement best describes how often and how long your science class meets?						
We meet every day for forty-five minutes to an hour.	95	71	1	18	52	29
We meet on alternate days for 80-90 minutes.	0	14	1	18	51	31
We meet every day for forty-five minutes, plus a longer lab period each week.	2	5	1	18	47	35
We have a flexible schedule depending on activities.	3	11	0	12	46	42
How often do you do assignments for science or take tests where you earn points for what you have written even if it is not completely correct?						
most of the time	38	36	1	18	53	29
some of the time	45	53	1	17	50	33
never	16	11	1	19	50	29
"I learn in school most of what I need to know to answer the MEA science & technology questions."						
strongly agree	4	8	2	26	45	27
agree	64	54	1	19	52	28
disagree	24	29	0	14	51	35
strongly disagree	8	9	0	9	47	43
How have you learned about Maine's *Learning Results*?						
from the science teacher who identifies the standards related to our course	24	25	1	19	51	29
from other teachers and the principal but not from science teachers	9	17	1	16	50	33
from newspaper or television	10	12	0	15	47	38
I have never heard of Maine's *Learning Results*.	57	46	1	17	52	30
"My knowledge of science and technology will be useful to me in my future work."						
strongly agree	22	31	1	25	52	22
agree	54	52	0	15	52	32
disagree	19	14	0	10	47	43
strongly disagree	5	4	0	7	40	53
Which courses do you plan to take before you graduate from high school?						
earth and space science and biology	25	23	0	13	54	33
the course(s) described above plus chemistry	15	19	1	19	52	28
the course(s) described above plus physics	21	27	1	30	48	21
a life science and physical science course	38	31	0	10	52	38
How well can you learn the science and technology that is taught in your school?						
very well	22	26	2	31	51	16
well	58	59	0	14	53	32
not too well	16	12	0	7	41	52
not well at all	5	3	0	4	37	59

MAINE EDUCATIONAL ASSESSMENT

SOCIAL STUDIES RESULTS

School:
District:
Grade: 8
Date: March 2000

PERFORMANCE LEVELS		STUDENTS AT EACH PERFORMANCE LEVEL				
		School		District		State
		N	%	N	%	%
Exceeds the Standards - The quality of a student's work at this level of proficiency exceeds the standards of performance as identified for Maine's *Learning Results* in social studies. The student demonstrates exemplary knowledge of content of major social studies concepts, consistently applies complex thinking skills, and communicates ideas clearly in all situations. Scaled scores (561-580).	1998-1999	5	2	5	2	1
	1999-2000	**7**	**3**	**7**	**3**	**2**
	2000-2001					
	Cumulative Average	6	3	6	3	2
Meets the Standards - The quality of a student's work at this level of proficiency meets the standards of performance as identified for Maine's *Learning Results* in social studies. The student demonstrates consistent knowledge of content of major social studies concepts, usually applies complex thinking skills, and communicates ideas clearly in most situations. Scaled scores (541-560).	1998-1999	56	25	56	25	24
	1999-2000	**59**	**24**	**59**	**24**	**27**
	2000-2001					
	Cumulative Average	58	25	58	25	26
Partially Meets the Standards - The quality of a student's work at this level of proficiency partially meets the standards of performance as identified for Maine's *Learning Results* in social studies. The student demonstrates some knowledge of content of major social studies concepts, inconsistently applies complex thinking skills, and communicates ideas clearly in some situations. Scaled scores (521-540).	1998-1999	120	54	120	54	55
	1999-2000	**129**	**53**	**129**	**53**	**50**
	2000-2001					
	Cumulative Average	125	54	125	54	53
Does Not Meet the Standards - The quality of a student's work at this level of proficiency does not meet the standards of performance as identified for Maine's *Learning Results* in social studies. The student demonstrates a limited knowledge of content of major social studies concepts, does not apply complex thinking skills, and communicates ideas clearly in few or no situations. Scaled scores (501-520).	1998-1999	43	19	43	19	20
	1999-2000	**48**	**20**	**48**	**20**	**21**
	2000-2001					
	Cumulative Average	46	20	46	20	21

Learning Results Content Standards	Number of Points Possible	Average Points Attained (Number and Percent)					
		School		District		State	
		N	%	N	%	N	%
Content	122	65.1	53	65.1	53	64.1	53
Application	72	26.5	37	26.5	37	26.4	37
Civics and Government (Standards A, B, C, and D)	49	22.2	45	22.2	45	21.4	44
Rights, Responsibilities, and Participation (Standard A)	16	8.7	54	8.7	54	8.1	51
Purpose, Types, and Fundamental Principles of Government and Constitutions (Standards B and C)	20	8.8	44	8.8	44	9.2	46
International Relations (Standard D)	13	4.7	36	4.7	36	4.1	32
History (Standards A, B, and C)	58	25.8	44	25.8	44	25.9	45
Chronology, Historical Knowledge, Concepts, and Patterns (Standards A and B)	38	16.4	43	16.4	43	16.8	44
Historical Inquiry, Analysis, and Interpretation (Standard C)	20	9.5	48	9.5	48	9.1	46
Geography (Standards A and B)	48	25.6	53	25.6	53	25.5	53
Skills and Tools (Standard A)	14	9.4	67	9.4	67	9.2	66
Human Interaction with Environments (Standard B)	34	16.3	48	16.3	48	16.3	48
Economics (Standards A, B, C, and D)	39	18.0	46	18.0	46	17.7	45
Personal and Consumer Economics (Standard A)	12	7.0	58	7.0	58	7.7	64
Economic Systems/Comparative Systems (Standards B and C)	19	7.8	41	7.8	41	6.9	36
International Trade and Global Interdependence (Standard D)	8	3.2	40	3.2	40	3.1	39

MAINE EDUCATIONAL ASSESSMENT

SOCIAL STUDIES RESULTS

(CONTINUED)

School:
District:
Grade: 8
Date: March 2000

Reporting Categories	School					State				
	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
GENDER										
boy	48	3	23	54	20	51	2	28	49	22
girl	52	2	27	53	19	49	2	26	51	20
INTERNET-CONNECTED COMPUTER IN HOME										
yes	74	3	26	54	17	72	2	30	50	18
no	24	2	22	51	25	27	1	19	51	29
I don't know.						1	1	14	43	43
QUANTITY OF TV ON A SCHOOL NIGHT										
none	4	0	30	40	30	6	4	35	43	18
less than one hour	18	2	37	42	19	25	2	31	48	18
one to two hours	51	3	25	57	15	43	2	29	51	18
more than two hours	27	3	14	55	28	27	1	19	51	29
H.O.T.S.										
students in a H.O.T.S. program						2	1	11	48	40
students completed program 2 years prior						1	1	13	47	39
TITLE 1 PROGRAM										
students currently served in math						2	0	7	43	49
students previously served in math						3	1	12	48	39
new students currently served in reading						1	0	2	41	58
new students previously served in reading	7	0	0	53	47	2	0	5	40	55
MIGRANT										
students eligible, not served						0	0	16	68	16
students eligible, served, not tutored						1	0	15	42	42
students eligible, served, tutored						1	0	10	50	40
STATE-APPROVED GIFTED/ TALENTED PROGRAM										
yes						4	13	67	18	2
no	100	3	24	53	19	96	1	25	51	22
ENGLISH LANGUAGE FLUENCY										
monolingual English	96	1	28	56	15	98	2	27	50	21
bilingual/English fluent						1	1	27	48	24
bilingual/limited-English proficient						1	1	6	61	31
IDENTIFIED DISABILITY										
yes	9	0	5	45	50	11	0	5	37	58
no	91	3	26	54	17	89	2	30	52	17
OPTIONAL SCHOOL/DISTRICT QUESTION										
A										
B										
C										
D										
E										

Questionnaire Items	Sch.	State				
	% Students in Each Category	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
How do you spend most of your class time in social studies?						
I work by myself.	36	22	2	24	49	26
I work in small groups.	6	16	1	19	53	28
I do some work by myself and some in small groups.	48	50	2	30	51	17
The whole class works together.	10	12	3	32	47	19
"I learn in school most of what I need to know to answer the MEA social studies questions."						
strongly agree	17	14	2	33	47	17
agree	63	57	2	28	51	19
disagree	17	23	2	24	51	24
strongly disagree	3	7	2	17	48	34
Did you have any experiences outside the classroom, such as fields trips, that helped you better understand your studying in social studies?						
Yes, we went on field trips.	22	22	2	29	49	20
Yes, we did projects in the community.	8	9	1	18	50	31
a combination of the options above	6	10	1	21	49	30
No, we did not go on field trips or do any projects.	63	59	2	29	51	18
Think about a research project that you did in social studies this year. What resources did you use?						
magazines, newspapers, books, and an encyclopedia	30	16	1	22	49	27
the Internet and/or personal interviews	4	11	1	17	49	33
a combination of the options above	44	62	2	31	51	16
I did not do any research projects in social studies.	22	11	1	24	46	29
How well can you learn the social studies that is taught in your school?						
very well	51	36	4	39	46	12
well	42	51	1	23	54	23
not too well	6	10	0	12	49	39
not well at all	1	3	0	6	47	47
How important is it for you to do well in social studies?						
very important	50	49	2	31	48	18
somewhat important	43	42	1	25	53	21
minimally important	5	7	1	17	49	33
not important	1	3	1	13	40	46
"My knowledge of social studies will be useful to me in my future work."						
strongly agree	22	18	2	32	46	19
agree	58	54	2	28	50	20
disagree	16	21	1	24	52	23
strongly disagree	4	7	1	17	51	31
How important is social studies compared to other courses or subjects that you are taking?						
very important	22	17	2	31	45	21
somewhat important	61	58	2	28	51	19
minimally important	15	19	1	23	52	23
not important	2	6	1	16	48	35

MAINE EDUCATIONAL ASSESSMENT

VISUAL & PERFORMING ARTS RESULTS

School:
District:
Grade: 8
Date: March 2000

PERFORMANCE LEVELS		STUDENTS AT EACH PERFORMANCE LEVEL				
		School		District		State
		N	%	N	%	%
Exceeds the Standards - The quality of a student's work at this level of proficiency exceeds the standards of performance as identified for Maine's *Learning Results* in visual and performing arts. The student demonstrates exemplary knowledge of content and application of skills of the visual and performing arts, including creative expression, cultural heritage, and criticism and aesthetics. Scaled scores (561-580).	1998-1999	3	1	3	1	1
	1999-2000	**11**	**5**	**11**	**5**	**5**
	2000-2001					
	Cumulative Average	7	3	7	3	3
Meets the Standards - The quality of a student's work at this level of proficiency meets the standards of performance as identified for Maine's *Learning Results* in visual and performing arts. The student demonstrates consistent knowledge of content and application of skills of the visual and performing arts, including creative expression, cultural heritage, and criticism and aesthetics. Scaled scores (541-560).	1998-1999	70	31	70	31	27
	1999-2000	**59**	**24**	**59**	**24**	**26**
	2000-2001					
	Cumulative Average	65	28	65	28	27
Partially Meets the Standards - The quality of a student's work at this level of proficiency partially meets the standards of performance as identified for Maine's *Learning Results* in visual and performing arts. The student demonstrates partial and/or inconsistent knowledge of content and application of skills of the visual and performing arts, including creative expression, cultural heritage, and criticism and aesthetics. Scaled scores (521-540).	1998-1999	120	53	120	53	53
	1999-2000	**87**	**36**	**87**	**36**	**35**
	2000-2001					
	Cumulative Average	104	45	104	45	44
Does Not Meet the Standards - The quality of a student's work at this level of proficiency does not meet the standards of performance as identified for Maine's *Learning Results* in visual and performing arts. The student demonstrates limited knowledge of content and application of skills of the visual and performing arts, including creative expression, cultural heritage, and criticism and aesthetics. Scaled scores (501-520).	1998-1999	33	15	33	15	19
	1999-2000	**85**	**35**	**85**	**35**	**34**
	2000-2001					
	Cumulative Average	59	25	59	25	27

Learning Results Content Standards	Number of Points Possible	Average Points Attained (Number and Percent)					
		School		District		State	
		N	%	N	%	N	%
Dance	22	11.0	50	11.0	50	10.4	47
Music	37	17.7	48	17.7	48	19.4	52
Theater	24	13.6	57	13.6	57	13.3	55
Visual Arts	37	21.1	57	21.1	57	20.7	56
Creative Expression (Standard A)	44	20.3	46	20.3	46	21.9	50
Cultural Heritage (Standard B)	43	23.6	55	23.6	55	23.3	54
Criticism and Aesthetics (Standard C)	33	19.4	59	19.4	59	18.7	57

MAINE EDUCATIONAL ASSESSMENT

VISUAL & PERFORMING ARTS RESULTS

(CONTINUED)

School:
District:
Grade: 8
Date: March 2000

Reporting Categories	School					State				
	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
GENDER										
boy	48	4	17	35	43	51	4	22	35	39
girl	52	5	31	37	27	49	6	30	35	29
INTERNET-CONNECTED COMPUTER IN HOME										
yes	74	5	28	32	35	72	6	28	35	31
no	24	3	15	44	37	27	3	21	34	42
I don't know.						1	1	14	30	55
QUANTITY OF TV ON A SCHOOL NIGHT										
none	4	10	30	30	30	6	9	33	31	27
less than one hour	18	5	28	35	33	25	6	29	34	31
one to two hours	51	5	27	38	31	43	5	28	36	31
more than two hours	27	3	17	34	46	27	3	19	35	43
H.O.T.S.										
students in a H.O.T.S. program						2	2	14	34	51
students completed program 2 years prior						1	2	19	33	47
TITLE 1 PROGRAM										
students currently served in math						2	1	8	35	55
students previously served in math						3	2	13	38	47
new students currently served in reading						1	0	5	34	61
new students previously served in reading	7	0	13	44	44	2	1	8	32	60
MIGRANT										
students eligible, not served						0	4	17	33	46
students eligible, served, not tutored						1	3	16	29	52
students eligible, served, tutored						1	0	14	24	63
STATE-APPROVED GIFTED/ TALENTED PROGRAM										
yes						4	22	50	20	8
no	100	5	24	36	35	96	4	25	36	35
ENGLISH LANGUAGE FLUENCY										
monolingual English	96	4	27	39	31	98	5	26	35	34
bilingual/English fluent						1	2	23	29	46
bilingual/limited-English proficient						1	4	14	31	52
IDENTIFIED DISABILITY										
yes	9	0	10	38	52	11	1	7	27	66
no	91	5	26	36	33	89	5	29	36	30
OPTIONAL SCHOOL/DISTRICT QUESTION										
A										
B										
C										
D										
E										

Questionnaire Items	Sch.	State				
	% Students in Each Category	% Students in Each Category	% Exceeds the Standards	% Meets the Standards	% Partially Meets the Standards	% Does Not Meet the Standards
How well can you learn the visual and performing arts that are taught in your school?						
very well	26	24	9	36	31	24
well	46	48	4	27	36	34
not too well	18	17	3	20	37	40
not well at all	9	12	1	17	38	44
"I learn in school what I need to know to answer the MEA visual and performing arts questions?						
strongly agree	8	11	8	30	31	31
agree	36	38	5	26	34	35
disagree	32	32	5	28	36	32
strongly disagree	24	20	3	22	38	36
What best describes your participation in music?						
I take or took a course at school.	71	62	6	30	36	28
I have not taken a course at school.	20	24	1	18	35	46
I am involved outside of school.	6	9	8	29	29	34
My school does not offer opportunities.	2	4	1	14	33	51
What best describes your participation in visual arts?						
I take or took a course at school.	53	50	7	31	34	28
I have not taken a course at school.	32	31	3	20	36	41
I am involved outside of school.	4	6	7	26	30	37
My school does not offer opportunities.	11	13	3	23	37	37
What best describes your participation in theater?						
I take or took a course at school.	8	16	7	33	34	26
I have not taken a course at school.	66	53	4	25	36	35
I am involved outside of school.	6	9	10	34	29	26
My school does not offer opportunities.	20	22	3	22	36	38
What best describes your participation in dance?						
I take or took a course at school.	5	8	3	24	33	40
I have not taken a course at school.	40	40	4	22	36	38
I am involved outside of school.	15	12	7	32	32	29
My school does not offer opportunities.	40	40	6	29	35	30
"My knowledge of visual and performing arts will be useful to me in my future work."						
strongly agree	25	18	8	32	32	28
agree	41	42	5	28	34	33
disagree	23	28	4	24	37	34
strongly disagree	11	13	2	17	38	42

MAINE EDUCATIONAL ASSESSMENT

Common Item Class Report

MATHEMATICS

GRADE 8

Code:
District:
School:
Class:
Date: March 2000
Group Size: 19 Page: 1 OF 1

Name	Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	18	19	20	21	23	24	25	26	27	28	29	33	35	37	Points Earned (50 Max. Points)	Scaled Score	Performance Level
	Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	SA	SA	SA	SA	CR	CR	MC	MC	MC	MC	MC	SA	CR	ER			
	Correct MC Response	A	D	B	D	D	C	A	C	A	A	D	B	A	B	A							D	C	C	B	B						
	Total Possible Points	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	4	4	1	1	1	1	1	2	4	8			
		+	C	+	+	+	+	D	+	+	B	+	+	+	D	C	0	0	1	0	0	0	+	+	+	+	+	2	1	2	21	524	P
		+	+	+	+	+	B	D	+	+	+	+	+	+	+	+	2	2	1	0	1	1	+	+	+	+	+	1	0	8	34	544	M
		+	+	+	C	+	+	C	A	D	C	A	D	+	A	C	2	0	0	0	1	1	B	D	+	A	A	0	4	5	20	524	P
		+	+	+	B	+	+	D	+	+	+	A	+	+	+	B	0	2	0	0	1	3	+	+	+	A	+	0	4	4	29	536	P
		+	C	D	B	C	+	B	A	D	C	A	+	C	C	D	0	0	0	0	1	2	B	+	A	+	+	0	0	1	10	508	D
		+	+	+	C	+	+	+	+	+	+	A	+	+	+	+	0	1	2	0	1	3	+	B	+	C	+	2	3	5	33	542	M
		+	C	C	C	+	+	B	+	D	B	A	A	B	A	C	0	1	0	0	3	2	B	B	A	C	+	1	1	3	16	518	D
		+	C	C	C	+	+	B	A	+	+	A	+	+	+	+	1	0	2	0	2	2	+	+	+	A	+	1	3	B	24	530	P
		+	+	+	+	+	+	B	+	+	D	A	D	+	C	+	0	2	2	0	2	3	+	+	+	C	D	0	4	2	28	536	P
		+	+	+	C	+	+	+	+	+	+	+	+	+	+	+	0	2	0	0	1	2	B	+	+	+	+	2	3	2	30	538	P
		+	B	+	C	+	+	D	A	+	C	A	+	B	C	+	0	1	1	0	1	3	B	+	+	C	+	0	2	3	21	524	P
		+	+	+	C	+	D	D	A	+	D	+	+	C	A	D	0	0	0	0	1	2	A	+	+	+	+	1	2	1	18	520	D
		+	B	+	C	+	B	C	A	+	B	A	A	B	A	+	1	1	0	0	1	1	C	+	D	C	+	0	1	2	14	514	D
		+	+	A	+	+	B	D	A	+	+	+	+	+	A	D	0	1	0	0	2	2	+	B	+	+	+	1	2	3	24	530	P
		+	C	C	+	C	+	B	A	+	D	A	+	C	+	B	2	1	0	0	2	3	B	+	A	D	+	0	0	1	17	518	D
		+	+	D	+	+	+	D	A	+	D	A	+	B	C	+	2	2	0	0	2	2	+	+	+	+	+	0	2	4	27	534	P
		+	+	C	+	+	+	+	A	+	B	B	+	B	+	+	2	0	2	0	1	2	+	+	+	A	+	2	2	2	27	534	P
		+	B	+	C	+	D	D	A	+	+	A	+	+	+	+	2	2	0	0	1	B	B	B	B	+	A	0	1	1	17	518	D
		+	C	+	B	+	+	D	A	B	B	A	A	+	A	C	2	0	0	0	3	1	+	D	+	A	+	2	1	3	20	524	P

Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	18	19	20	21	23	24	25	26	27	28	29	33	35	37			
Percent Correct/Avg. Score: Class	100	56	67	39	89	72	17	33	83	39	28	78	61	44	56	0.9	0.9	0.6	0.0	1.3	1.8	56	72	78	44	83	0.8	1.9	2.7			
Percent Correct/Avg. Score: School	89	61	66	37	76	74	21	29	74	34	26	55	58	47	53	0.8	0.9	0.5	0.0	1.4	1.4	45	74	66	37	71	0.5	1.7	2.0			
Percent Correct/Avg. Score: District	76	59	57	42	67	66	22	26	68	34	46	55	57	51	49	0.6	0.9	0.4	0.1	1.1	1.4	52	66	55	36	57	0.5	1.8	2.0			
Percent Correct/Avg. Score: State	82	67	63	54	68	72	37	39	72	38	56	57	62	55	55	0.8	1.1	0.5	0.3	1.4	1.6	58	66	58	39	60	0.8	1.8	2.3			

MAINE EDUCATIONAL ASSESSMENT

Common Item Class Report

SCIENCE & TECHNOLOGY

GRADE 8

Code:
District:
School:
Class:
Date: March 2000
Group Size: 19 Page: 1 OF 1

Name	Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	16	18	19a	19bc	20	21	22	23	24	25	26	31	32	33	34	35	36	Points Earned (50 Max. Points)	Scaled Score	Performance Level
	Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	SA	CR	ER	ER	MC	MC	MC	MC	MC	MC	MC	SA	SA	SA	SA	CR	CR			
	Correct MC Response	B	A	A	D	B	C	D	B	D	A	D	C	A					D	B	A	D	C	C	B									
	Total Possible Points	1	1	1	1	1	1	1	1	1	1	1	1	1	2	4	4	4	1	1	1	1	1	1	1	2	2	2	2	4	4			
[redacted]		D	B	+	+	+	B	+	C	+	D	+	B	C	0	0	2	1	B	+	C	+	+	+	+	1	2	2	2	1	2	24	530	P
		+	+	+	+	+	+	+	C	+	+	+	+	+	2	2	4	2	+	+	+	+	+	+	+	1	2	2	2	1	2	39	556	M
		C	D	+	B	+	B	A	D	+	D	C	+	+	0	3	0	1	+	+	C	B	+	+	D	1	0	0	2	4	2	22	526	P
		+	D	+	+	D	+	C	+	A	+	B	+	+	0	2	2	0	+	+	+	C	+	+	+	0	2	2	2	2	2	28	536	P
		D	+	B	B	A	+	A	D	A	C	+	B	D	0	0	0	0	B	C	+	+	B	A	D	1	1	0	0	0	1	8	504	D
		+	+	+	+	+	+	+	+	A	+	+	D	+	2	3	4	3	+	+	+	A	+	+	+	1	2	1	2	3	3	41	560	M
		+	D	+	A	+	+	B	A	A	D	+	A	+	0	1	3	0	B	+	+	B	B	A	A	1	0	1	0	1	1	16	516	D
		+	C	+	+	D	+	A	+	A	D	C	B	+	0	2	2	0	+	+	+	+	+	+	C	2	2	2	2	3	2	29	536	P
		+	+	+	B	+	+	C	C	A	+	+	+	+	2	2	2	1	+	+	+	+	+	+	+	1	0	2	2	1	2	31	540	P
		+	+	+	+	+	+	B	+	+	+	+	A	+	0	2	0	0	+	+	C	A	+	+	+	1	0	2	2	2	2	27	534	P
		+	D	B	+	A	D	+	+	+	C	+	A	+	0	2	1	1	+	C	+	+	+	A	+	1	0	1	2	2	0	22	526	P
		+	+	+	+	+	+	+	+	A	C	+	D	D	2	2	3	1	+	C	+	C	+	+	+	1	1	2	0	1	2	29	536	P
		+	D	+	C	+	+	C	+	+	+	+	A	D	0	0	2	1	+	+	C	B	+	A	+	1	0	0	2	1	2	21	524	P
		+	+	+	+	+	+	+	C	A	B	+	+	D	2	1	1	0	+	+	+	A	+	+	D	1	0	0	2	1	2	24	530	P
		+	D	+	+	D	+	B	+	A	D	+		D	0	3	3	1	+	+	C	A	B	+	+	1	2	1	2	1	1	25	530	P
		C	D	+	+	+	+	B	+	A	+	C	+	+	0	2	1	1	B	C	+	+	+	+	A	1	0	2	2	2	2	25	530	P
		+	+	+	+	A	B	+	+	+	C	B	D	+	2	2	0	1	+	+	+	+	+	+	+	1	2	2	2	3	2	32	542	M
		D	B	B	+	+	+	B	C	+	C	+	A	D	0	1	3	0	A	C	+	A	+	+	D	1	B	1	2	1	1	18	520	D
		+	+	+	+	+	+	+	+	A	C	C	+	+	0	0	2	2	+	+	B	+	+	+	+	1	2	2	2	1	2	30	538	P

Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	16	18	19a	19bc	20	21	22	23	24	25	26	31	32	33	34	35	36
Percent Correct/Avg. Score: Class	72	50	83	78	67	78	44	61	44	39	67	39	61	0.7	1.6	1.8	0.9	78	72	67	50	89	83	67	1.0	1.0	1.3	1.8	1.7	1.8
Percent Correct/Avg. Score: School	63	53	61	74	50	63	53	61	42	39	74	29	58	0.5	1.2	1.7	0.7	76	66	45	39	76	79	55	0.9	0.8	1.1	1.5	1.5	1.5
Percent Correct/Avg. Score: District	51	50	57	64	53	61	53	56	46	44	75	35	49	0.4	0.8	1.6	0.7	66	63	41	39	84	67	51	0.9	0.8	1.0	0.9	1.2	1.2
Percent Correct/Avg. Score: State	54	65	64	62	49	71	57	62	49	54	71	45	51	0.5	0.9	1.9	0.9	73	63	51	46	87	65	57	1.0	1.2	1.1	1.1	1.3	1.4

MAINE EDUCATIONAL ASSESSMENT

Common Item Class Report

SOCIAL STUDIES

GRADE 8

Code:
District:
School:
Class:
Date: March 2000
Group Size: 19 Page: 1 OF 1

Name	Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	16	18	19ab	19c	20	21	22	23	24	25	26	31	32	33	34	35	36	Points Earned (50 Max. Points)	Scaled Score	Performance Level
	Item Type	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	MC	SA	CR	ER	ER	MC	MC	MC	MC	MC	MC	MC	SA	SA	SA	SA	CR	CR			
	Correct MC Response	C	B	D	C	B	C	A	D	B	A	C	A	D					D	B	A	B	C	A	C									
	Total Possible Points	1	1	1	1	1	1	1	1	1	1	1	1	1	2	4	4	4	1	1	1	1	1	1	1	2	2	2	2	4	4			
		+	A	+	+	+	+	+	+	+	D	+	C	+	2	2	2	2	+	C	+	+	B	+	+	1	2	0	0	0	1	27	536	P
		+	+	B	+	+	+	+	+	+	+	+	+	+	2	1	4	4	+	+	+	A	+	+	+	2	2	2	1	1	2	39	560	M
		+	+	+	B	+	+	+	+	+	B	+	D	+	1	1	3	1	+	C	D	+	+	+	+	1	1	2	1	0	1	27	536	P
		+	+	+	+	+	+	+	B	+	+	+	+	C	2	3	4	4	+	+	D	+	+	+	+	1	2	0	0	0	3	36	554	M
		+	A	B	+	+	+	D	+	A	C	+	B	B	1	0	1	0	B	C	D	C	A	+	+	1	2	0	0	0	1	14	516	D
		D	+	+	+	+	+	+	A	+	+	+	B	+	2	3	3	4	+	+	+	+	+	+	+	1	2	2	2	1	3	40	562	E
		B	A	C	D	+	+	B	B	A	+	B	D	A	0	2	1	2	+	+	+	+	B	+	D	1	1	0	2	0	2	19	524	P
		+	+	B	+	+	+	+	A	+	+	B	+	+	2	4	3	3	+	+	+	+	+	+	+	1	2	2	1	1	2	38	558	M
		+	+	+	+	+	+	+	+	+	+	+	B	+	1	1	3	3	+	A	+	+	A	+	B	1	1	B	2	1	2	31	544	M
		+	+	+	+	+	+	+	+	+	+	+	B	+	2	3	3	3	+	+	+	+	+	+	+	1	2	2	2	1	3	41	564	E
		+	A	+	+	+	+	+	B	D	D	+	D	+	0	3	2	0	+	C	C	+	+	+	+	2	2	0	2	1	1	26	534	P
		+	C	+	+	C	+	+	+	C	+	+	+	+	1	2	3	1	+	A	+	+	+	+	D	1	2	0	1	0	2	28	538	P
		+	+	A	D	C	+	C	B	+	B	+	B	C	0	2	3	1	+	C	+	+	+	B	+	1	1	0	0	0	2	20	526	P
		D	A	A	+	+	+	+	C	+	B	+	C	+	1	3	1	2	B	C	B	+	A	+	D	0	2	0	1	0	2	21	526	P
		+	A	B	+	+	+	+	B	A	+	+	B	B	2	2	0	0	+	+	C	+	+	+	D	1	2	B	2	0	2	23	530	P
		+	+	+	+	+	+	+	+	+	B	+	+	+	2	3	3	3	+	+	+	+	B	+	+	2	2	2	1	1	2	39	560	M
		+	+	+	+	+	+	+	+	+	B	+	+	+	2	2	3	3	+	+	+	+	+	+	+	2	2	2	2	0	4	41	564	E
		+	+	A	+	+	+	C	A	+	D	+	+	+	1	1	2	0	+	C	+	+	B	C	+	2	1	2	2	1	2	27	536	P
		+	+	B	+	+	+	+	B	+	+	+	+	+	2	2	2	1	+	+	D	+	+	+	+	2	2	B	2	0	2	32	546	M
Item Number		1	2	3	4	5	6	7	8	9	10	11	12	13	16	18	19ab	19c	20	21	22	23	24	25	26	31	32	33	34	35	36			
Percent Correct/Avg. Score: Class		89	67	56	89	89	100	83	50	78	50	94	44	78	1.4	2.1	2.5	1.9	89	50	61	89	67	89	78	1.3	1.8	0.9	1.2	0.4	2.1			
Percent Correct/Avg. Score: School		87	58	37	89	84	89	74	53	68	37	74	37	63	1.4	1.6	2.1	1.4	89	45	63	82	58	87	71	1.0	1.7	0.6	1.2	0.2	1.8			
Percent Correct/Avg. Score: District		82	52	41	72	78	90	61	44	50	31	72	30	61	1.2	1.3	1.5	1.0	87	41	64	60	57	86	55	1.0	1.0	0.5	1.0	0.2	1.3			
Percent Correct/Avg. Score: State		84	52	35	73	75	90	66	51	57	32	76	30	70	1.3	1.6	1.8	1.3	85	54	72	65	59	91	62	1.0	1.0	0.8	1.1	0.2	1.6			